Metabolic Engineering of Yeast for Xylose Uptake and Fermentation

by

Hang Zhou

B.E. Chemical Engineering, Tsinghua University (2002)

M.S. Chemical Engineering, Tsinghua University (2005)

M.S. Chemical Engineering Practice, Massachusetts Institute of Technology (2009)

Submitted to the Department of Chemical Engineering in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Chemical Engineering
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JANUARY 2011
[February 2011]

Signature of author.....
Department of Chemical Engineering
January 13, 2011

Certified by.....
Gregory N. Stephanopoulos
W.H. Dow Professor of Chemical Engineering and Biotechnology
Thesis Advisor

Accepted by.....
William M. Deen
Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Metabolic Engineering of Yeast for Xylose Uptake and Fermentation

by

Hang Zhou

Submitted to the Department of Chemical Engineering on January 13, 2011
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Chemical Engineering

Abstract

Xylose is the major pentose and the second most abundant sugar in lignocellulosic feedstocks. Therefore, the efficient utilization of xylose is required for the cost-effective bioconversion of lignocellulose. Rational and combinatorial metabolic engineering approaches coupled with transcriptomic and proteomic studies have been extensively exploited on the yeast *Saccharomyces cerevisiae* for improved xylose utilization. However, the resulting strains remain inapplicable for industrial ethanolic fermentation, although basic engineering approaches have been established and targets have been identified for further modification.

In this study, we started with the rational genetic engineering of a laboratory *S. cerevisiae* strain to express the xylose metabolic pathways, including the xylose reductase/xylitol dehydrogenase (XR/XDH) and xylose isomerase (XI) pathways. The xylulokinase (XK) and the non-oxidative pentose phosphate pathway (PPP) were also overexpressed to facilitate pentose assimilation. The resulting strains, H131-XYL123 and H131-XYLA31, exhibited slow but significant aerobic growth (μ_{max}=0.031±0.022 h^{-1} and 0.081±0.052 h^{-1}, respectively), establishing a baseline for further advancement.

These engineered strains mentioned above were then used to initiate a three-stage evolutionary engineering, through aerobic and oxygen-limited sequential batch cultivation followed by xylose-limited anaerobic chemostat cultivation. We aborted the development of the XR/XDH-based strain after the first stage of aerobic evolution, owing to the intrinsic cofactor imbalance of the two enzymes. In contrast, continuous improvement was observed during adaption of the XI-based H131-XYLA31, when strains were isolated from the evolved populations periodically and evaluated in terms of growth and fermentation. The final isolated strain, H131E8-XYLA31, displayed a significantly increased anaerobic growth rate (0.120±0.004 h^{-1}) and xylose consumption rate (0.916 $g \cdot g^{-1} h^{-1}$) compared to its parent strain.

Upon successful evolutionary engineering, the H131E8-XYLA31 was further modified by complementing the auxotrophic markers *arg4* and *leu2*, resulting in H153E10-XYLA31 with greatly boosted aerobic growth. Moreover, adding the anaerobic growth factors ergosterol and Tween 80 to the medium enabled a maximum anaerobic growth rate of 0.199 h^{-1} and a specific xylose consumption rate of 1.647$g \cdot g^{-1} h^{-1}$ in batch fermentation, 65% and 37% higher than those of the best reported xylose-

fermenting strain RWB 218, respectively. In chemostat cultivation, the strains exhibit specific performance more than 50% better than in the batch cultivation, implying the potential for further improvement of the strains in extended evolution.

In order to identify the genotypes responsible for quick xylose utilization in the evolutionarily engineered strain, an inverse inverse metabolic engineering approach was applied to the evolved strain, based on functional complementation of the evolved H131E5-XYLA31 genomic library to an unevolved background strain. A high-throughput micro-fluidic screening method was used to screen the library, revealing the tandem duplication of *XYLA* in the H131E5-XYLA31 genome. The structure of *XYLA* integration, coupled with qPCR, DNA/RNA blotting, and enzyme activity assay results, suggests that the high expression level of XI is a major recombination event during the evolution and is necessary for efficient xylose assimilation. However, high XI expression does not necessarily imply the tandem multi-copy integration of *XYLA* into the genome. The *XYLA* expansion could originate from different initial constructions, subject to recombination/rearrangement, and produce different configurations.

The effects of other metabolic engineering targets were also investigated, including other heterologous XIs, XK (*XYL3*), aldose reductase (*GRE3*), NADP-dependent G3P dehydrogenase (*gapN*), and heterologous xylose transporters. Most results verified the previous hypothesis, leaving room for future strain improvements.

In this study, we successfully applied rational and combinatorial metabolic engineering approaches for both constructing rapid xylose-fermenting strains and identifying novel genetic characteristics. We demonstrated the use of evolutionary engineering to achieve superior strains, as well as the inverse metabolic engineering approach based on micro-fluidic screening for both strain improvement and genotype identification. As such, the work of this thesis serves as a practical milestone in lignocellulose conversions. The constructed strains can be used as hosts for further advancement, and the metabolic engineering techniques have proven to be effective tools for future strain evolution.

Thesis Supervisor: Gregory N. Stephanopoulos
W.H. Dow Professor of Chemical Engineering and Biotechnology

Acknowledgements

I would like to thank my thesis advisor, Prof. Greg Stephanopoulos, for his guidance and encouragement, for his patience when I struggled with my little bugs, and for the freedom he granted me to pursue my own ideas. My entire thesis committee also provided essential ideas and critical feedback. Thanks to Professor Gerry Fink and Narendra Maheshri for insights in yeast research. Thanks to Professor Daniel Wang for willingness to share both his expertise in fermentation and his laboratory resources. Professor Graham Walker, although not technically on my committee, offered help in searching for sugar transporters.

I have been lucky to work closely with many brilliant and dedicated researchers. Thanks to Yong-Su Jin and Hal Alper for being good teachers helping me to initiate my own project. Thanks to Benjamin Wang for working together with me in the library screening, and Chapter 4 was written in close collaboration with Ben. Thanks to Jingsheng Cheng, for his help in those exhausting biochemical and fermentation studies. Thanks to Kang Zhou for his expertise in qPCR. To Simon Carlsen, Mitchell Tai, Jose Avalos and Felix Lam, for sharing their experiences and advice in yeast research. To Keith Tyo, Christine Santos, and Adel Ghaderi for patiently reminding me to label all my bottles. To Rosangela dos Santos for her continuous administrative help. Also, many thanks go to other members of the Stephanopoulos lab, especially Curt Fischer, Jason Walther, Daniel Klein, Ajikumar Parayil, Hussain Abidi, Vikram Yadav, Tom Wasylenko, and so many others, for sharing the great times with me in the lab.

My gratitude goes to Huimin Yu, Hui Luo, Wenhai Xiao, Xiaobing Huang, Wen-chi Yang, Peipei Han, Claire Li and many others, for making my lunch time in a relaxing Chinese speaking community. Thanks to all my friends at MIT and elsewhere, too many to list here, for making my days at MIT and in US very enjoyable.

An extra sincere thank goes to my parents, Jinkun and Xiufang, who never stop their unconditional love and support throughout my entire life, and encourage me to always think independently, practice boldly and strive for excellence. I'm so happy to welcome my lovely boy Alex, who constantly surprises me by showing his new skills and becomes an extra reason that I enjoy every single day of my life. And finally, I would like to express my most hearty gratitude to my wife, Wei Guo, for her love, comprehension and encouragement, and it is to you that I dedicate this thesis.

Table of Contents

List of Figures

List of Tables

CHAPTER 1. **Introduction**

1.1 Research background and motivation

Developments in petrochemistry have greatly boosted industrial and societal progress in the 20th century. The ever-increasing rate of exploration and exploitation of oil reserves has, for a long time, supplied an abundant and cheap raw material for transport fuels as well as the production of bulk and fine chemicals. However, in the past decade, several factors such as the increasing demand for fossil resources, the short- and long-term uncertainties generated by geopolitical issues, and rapid increase of the CO_2 levels in the Earth's atmosphere and oceans which is related to the use of fossil resources and affects the global climate via the well-publicized "greenhouse effect"[1], are contributing to create an urgent need for alternatives to an economy that predominantly depends on fossil resources. Plant biomass offers an attractive alternative, and it would be extremely attractive to have cost-effective, sustainable means of producing transport fuels from plant biomass. One of the most promising processes in this respect is the production of fuel ethanol. Ethanol can be blended with conventional fuels or used alone, and modifications to conventional car motors are either limited or not required. Although CO_2 is emitted during fermentation, the net effect is offset by the uptake of carbon gases by the plants grown to produce ethanol. When compared to gasoline, ethanol releases less greenhouse gases [2].

Fuel ethanol production from plant biomass hydrolysates by *Saccharomyces cerevisiae* is of great economic significance. Wild-type *S. cerevisiae* strains readily

ferment glucose, mannose and fructose via the Embden–Meyerhof pathway of glycolysis, while galactose is fermented via the Leloir pathway [3]. However, consensus has been reached in the past decades that, in the long run, cost-effective production of ethanol from plant biomass should not only be based on the readily fermentable starch and sucrose fractions of plant carbohydrates, but also on the much more resistant lignocellulosic fractions. Lignocellulosic crop residues comprise more than half of the world's agricultural phytomass [4] and significant fractions of the total can be recovered without competing with other uses [5]. Xylose constitutes about 10-20% of the total dry weight in woody common agricultural lignocellulosic feedstocks (Table 1-1) [3].

Table 1-1 Major sugars of common agricultural lignocellulosic feedstocks

	Corn stover	Wheat straw	Bagasse	Cotton gin	Sugar beet pulp	Switch grass
Carbohydrate (%)						
Glucose	34.6	32.6	39.0	37.1	24.1	31.0
Mannose	0.4	0.3	0.4	1.1	4.6	0.2
Galactose	1.0	0.8	0.5	2.4	0.9	0.9
Xylose	19.3	19.2	22.1	9.4	18.2	0.4
Arabinose	2.5	2.4	2.1	2.3	1.5	2.8
Uronic acids	3.2	2.2	2.2	NA	20.7	1.2
Non-carbohydrate (%)						
Lignin	17.7	16.9	23.1	28.8	1.5	17.6
Extractives	7.7	13.0	3.8	7.7	NA	17.0
Ash	10.4	10.2	3.7	10.5	8.2	5.8

One source of xylose is the sulfite-pulping of hardwood [6]. Depending on the substrate and reaction conditions, dilute acid pretreatments of lignocellulosic residues can recover 80-95% of the xylose from the feedstock [7]. However, the efficient utilization of xylose presents a technical barrier to the commercial bioconversion, and therefore a major challenge in metabolic engineering.

Some yeasts, like *Candida utilis*, will grow on xylose but do not produce ethanol. In the 1980s, intensive screening efforts rapidly revealed that some yeasts, such as *Pachysolen tannophilus*, *Pichia stipitis, Candida shehatae* [8] and *Candida tropicalis* [9],

can convert xylose to ethanol directly under aerobic or oxygen-limiting conditions Many improvements have been made in the genetic engineering of yeasts and bacteria for the fermentation of xylose to ethanol. However, the bioconversion of pentoses to ethanol still presents a considerable economic and technical challenge [7].

The cost-effective production of ethanol requires several key factors:

1. The high-yield, high-rate fermentation of biomass hydrolysates to ethanol;
2. The tolerance to numerous compounds in the plant hydrolysates that inhibit microbial growth [10], as well as to high concentrations of ethanol;
3. The ability to grow under strictly anaerobic conditions. Overaeration leads to increased respiration and suboptimal ethanol yields, while an accurate level of oxygen in large-scale processes with viscous two-phase feedstocks is not only economically undesired, but also virtually impossible [11];
4. High resistance to contaminations. It is almost impossible to maintain aseptic conditions in industrial-scale ethanol production. In general, the combined use of low pH and high ethanol concentrations suffices to keep contaminations at bay in large-scale yeast-based ethanol production processes.

Based on its widespread, large-scale application for bioethanol production from hexoses, *S. cerevisiae* appears to be the most promising metabolic-engineering platform for bioethanol production from plant hydrolysates. Although native strains of *S. cerevisiae* do not use xylose as a carbon source, recent metabolic and evolutionary engineering studies on *S. cerevisiae* strains have greatly improved fermentation of this pentose.

Table 1-2 Characteristics of xylose-fermenting *S. cerevisiae* strains

Strain	Strain description	Fermentation condition	Yield g/g		$q_{Ethanol}$	q_{Xylose}	μ_{max}
			Ethanol	Xylitol	g/g•h	g/g•h	h^{-1}
1400 (pLNH32)[12]	*XYL1, XYL2, XKS1*	Fermentative batch, YPEX	0.30	0.08	0.36 g/L·h	1.07 g/L·h	0.29
424A(LNH-ST)[13]	*XYL1, XYL2, XKS1*	Fermentative batch	0.43	0.10	N/A	N/A	N/A
TMB 3255[14]	*XYL1, XYL2, XKS1, Δzwf1*	Oxygen limited, batch, synthetic medium	0.41	0.05	0.01	0.02	0.05
TMB 3271[15]	*XYL1*^{M}(K270M), *XYL2, XKS1*	Oxygen limited, batch, synthetic medium	0.31	0.09	0.07	0.226	0.06
TMB3057[16]	*XYL1, XYL2, XKS1, TAL1, TKL1, RPE1, RKI1, Δgre3*	Anaerobic batch, synthetic medium	0.33	0.33	0.04	0.13	0.16*
TMB3066[16]	*Piromyces XYLA, XKS1, TAL1, TKL1, RPE1, RKI1, Δgre3*	Anaerobic batch, synthetic medium	0.43	0.04	0.02	0.05	0.02*
FPL-YSX3C[17]	*XYL1, XYL2, XYL3*	Aerobic batch, minimal medium	0.19	0.20	0.019	0.10	0.058
FPL-YSX3 TAL1M[18]	*XYL1, XYL2, XYL3, TAL1*	Aerobic batch, minimal medium	0.21	0.31	0.033	0.16	0.123
RW 202-AFX[19]	*Piromyces XYLA* evolved isolate	Anaerobic batch, synthetic medium	0.42	0.021	0.14	0.34	0.03
RWB 217[20]	*Piromyces XYLA, XKS1, TAL1, TKL1, RPE1, RKI1, Δgre3*	Anaerobic batch, synthetic medium	0.43	0.003	0.46	1.06	0.09
RWB 218[3, 21]	Selection of RWB 217	Anaerobic batch, synthetic medium	0.41	0.001	0.49	1.2	0.12
INVSc1/pRS406XKS/ pILSUT1/pWOXYLA[22]	*Orpinomyces XYLA, XKS1, SUT1*	Aerobic batch	0.39	0.08	0.043 g/L·h	0.11 g/L·h	0.025
ADAP8[23]	*Orpinomyces XYLA XKS1, PsSUT1,* evolved isolate	Aerobic batch, complex medium	0.48	N/A	0.038	0.079	0.04
BWY10Xyl[24]	*Clostridium phytofermentans XYLA* in industrial strain	Anaerobic batch, synthetic medium	0.43	0.18	0.03	0.065	0.04*

* growth on D-xylose medium was tested under aerobic conditions

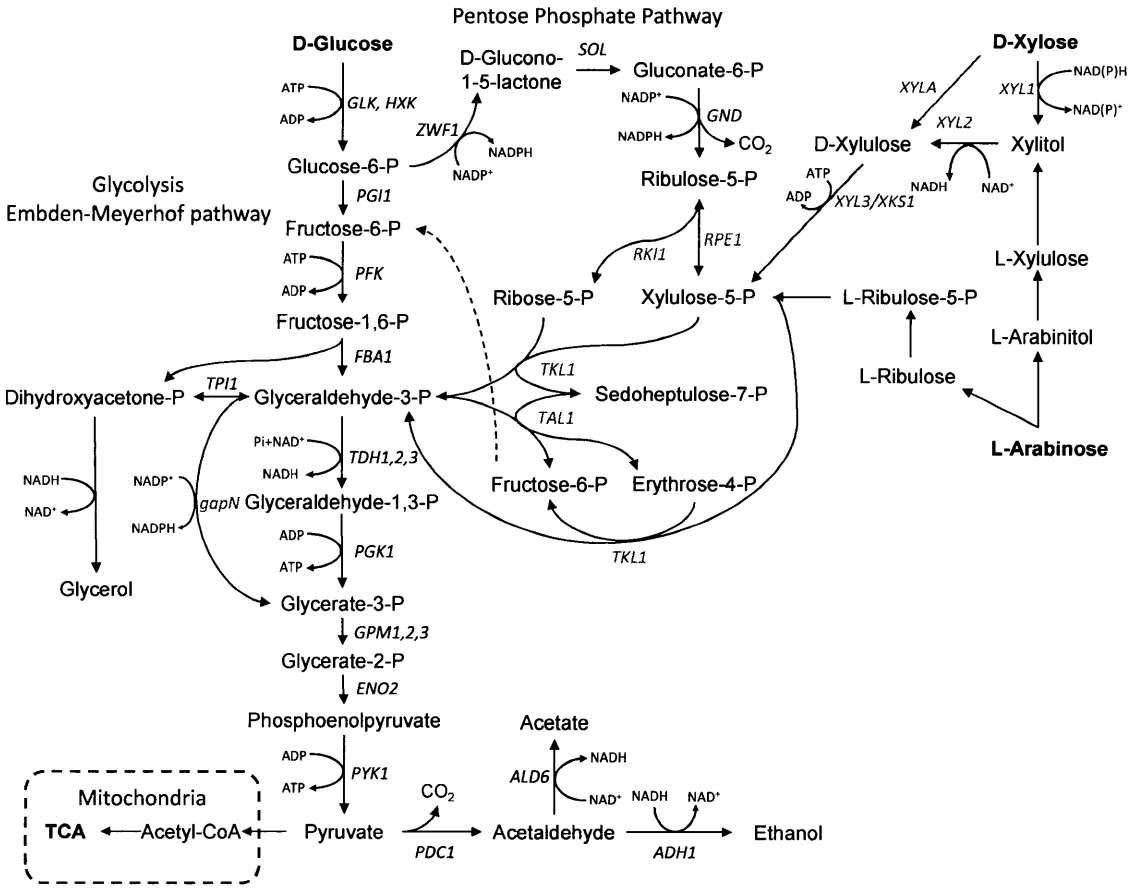


Figure 1-1. Metabolic network of xylose metabolism in the recombinant *S. cerevisiae*

1.2 Xylose reductase (XR) and xylitol dehydrogenase (XDH)

The assimilation of xylose requires two oxidoreductase steps, which are not available in *S. cerevisiae*. One of the best xylose-fermenting yeasts, *P. stipitis,* has served as the source of genes for engineering xylose metabolism in *S. cerevisiae.* The *XYL1* and *XYL2* (Figure 1-1) genes coding for xylose reductase (XR) and xylitol dehydrogenase (XDH) from *P. stipitis* have been expressed in *S. cerevisiae* [12, 25-27]. The resulting transformants can use xylose oxidatively and produce xylitol without the addition of a co-metabolizable carbon source, but only produce significant amounts of ethanol in low yield. The xylitol accumulation is mainly due to a cofactor imbalance, caused by the

different requirements for XR and XDH activities. XR from *P. stipitis* has a high affinity for NADPH even though it can use NADH as a cofactor. However, *P. stipitis* XDH only uses NAD^+ as a cofactor [28]. This difference results in cofactor imbalance in the cytosol during xylose metabolism, since no route is available for their regeneration. Respiration can alleviate the redox imbalance engineered *S. cerevisiae*. Under aerobic conditions the NADH can be reoxidized via the respiratory chain with molecular oxygen, and therefore, for optimal ethanol production, low aeration rates are required [29, 30]. But the exact aeration may be too costly to achieve in industrial-scale fermentation. To alleviate the redox imbalance and reduce the xylitol production, various attempts have been made to alter the cofactor specificity for XR and/or XDH [15, 31].

P. stipitis XR has about a twofold higher affinity for NADPH than for NADH [32, 33] and the intracellular ratio of NADPH/NADH is about 3 [34], so both facts favor the use of NADPH by XR. Watanabe et al. reported a mutated NADH-preferring PsXR (R276H), which was improved 52-fold in the ratio of NADH/NADPH in catalytic efficiency [(k_{cat}/K_m with NADH)/(k_{cat}/K_m with NADPH)][34, 35]. The K_m for NADH dropped by 45% while the K_m for NADPH went up more than 1000-fold, the K_m for xylose doubled and the K_{cat} with NADH stayed about the same. *S. cerevisiae* expressing PsXR R276H mutant showed a 20% increase in ethanol and a 52% decrease in xylitol production. Combinatorial approaches such as active site saturation (CASTing) is another empirical method for the protein engineering [36]. Selected amino acid residues were targeted with saturating mutagenesis in several successive rounds with high-throughput screening in each round to identify the best mutant. The best *P. stipitis XYL1* mutant, 2-2C12(K270S, N272P, S271G, and R276F) showed a NADH/NADPH activity ratio of

about 13, and a 73% K_{cat} with NADH. However, the K_m for NADH rose about 7-fold, but growth and fermentation characteristics were not reported [32]. A *Candida tenuis XYL1* K274R, N276D double mutant decreased xylitol production 52% while improving the ethanol yield 42% [37].

The cofactor imbalance also has been addressed by altering the preference of XDH from NAD^+ to $NADP^+$, albeit the reported $NAD^+/NADP^+$ ratio of 4.9 [34] favors NAD^+ usage by XDH. A double mutant of *Gluconobacter oxydans* XDH showed completely reversed specificity for $NADP^+$ and K_m for xylitol increased from 13.7 to 100 mM [38]. By substituting an XDH containing four mutations, xylitol production decreased 86% and ethanol production increased 41% [39]; increased specific xylose consumption and specific ethanol production rates in hydrolysate were also observed [40].

Fusion of XR and XDH has also been used in an attempt to get around cofactor imbalance. This resulted in a low ethanol yield, but the strain was capable of growth and ethanol production from xylose [41].

High levels of XR and XDH activity are important during xylose utilization [42, 43] and a higher level of XDH than XR reduces xylitol formation [39]. However, anaerobic xylose fermentation by engineered strains was inevitably accompanied by considerable xylitol production (Table 1-2) [3, 15].

1.3 Xylose isomerase (XI)

Functional expression of a heterologous xylose isomerase in *S. cerevisiae*, which converts xylose to xylulose directly (Figure 1-1), is another logical approach to engineer this yeast to ferment xylose, bypassing the cofactor imbalance problem in XR-XDH

pathway. However, attempts to introduce a xylose isomerase in *S. cerevisiae* were unsuccessful[3, 44-46] until 2003, when the first fungal xylose isomerase that could be functionally expressed in *S. cerevisiae* at high levels was discovered, which is encoded by the *XYLA* gene of the anaerobic fungus *Piromyces* sp E2 [47]. The introduction of the *XYLA* resulted in a strain (RWB 202-AFX) capable of anaerobic growth on xylose producing mainly ethanol, CO_2, glycerol and biomass and notably little xylitol (Table 1-2) [19].

XI from another anaerobic fungus *Orpinomyces*, which shows 94.3% amino acid sequence homology to *Piromyces* XI, have also been successfully expressed in *S. cerevisiae* [22]. Further evolutionary engineering increased the growth rate and ethanol yield of this strain but they are still not comparable to those of the evolved strain with *Piromyces* XI [23].

The cloning and successful expression of *Clostridium phytofermentans* XI in *S. cerevisiae* was also reported recently [24]. After a very short period of optimization, the industrial strain BWY10Xyl exhibited noticeable growth and fermentation on xylose, with an unusual high xylitol yield of 0.18 g/g of D-xylose consumed (Table 1-2). As an advantage, the XI from *C. phytofermentans* is significantly less susceptible to inhibition by xylitol (K_i=14.51±1.08 mM) than is the XI from the *Piromyces* strain (K_i=4.67±1.77 mM).

The XI pathway and the XR/XDH pathway were compared in the same background [16]. *S. cerevisiae* engineered with the XR/XDH pathway (TMB3057 in Table 1-2) had much higher aerobic growth and anaerobic xylose consumption rates than the XI strain (TMB3066 in Table 1-2), but the ethanol yield was lower due to the cofactor

imbalance. In hydrolysates, the XR/XDH pathway may have an advantage by detoxifying 5-hydroxymethyl-furfural (HMF) [48].

1.4 Xylulokinase (XK)

To obtain a higher specific rate of ethanol production, it is helpful to overexpress genes involved in the conversion of xylose into intermediates of glycolysis [17, 18, 20, 49, 50], including xylulokinase and the Pentose Phosphate Pathway (PPP).

The *S. cerevisiae* genome contains the gene *XKS1* coding for XK [51, 52], but the XK activity in wild-type *S. cerevisiae* is too low to support ethanolic xylose fermentation in strains engineered with a xylose pathway [30, 51, 53]. It is only when additional copies of *XKS1* are expressed that recombinant xylose-utilizing *S. cerevisiae* produce ethanol from xylose [54]. However, nonphysiological or unregulated kinase activity may cause a metabolic disorder [55]. It was indeed experimentally demonstrated that only fine-tuned overexpression of *XKS1* in *S. cerevisiae* led to improved xylose fermentation to ethanol [17, 56]. *XYL3* from *P. stipitis* was reported to have higher xylulokinase activity than endogenous *XKS1* [17], and also have high specific activity on xylulose than xylitol, avoiding any possible toxic xylitol-5P production[57].

1.5 Pentose Phosphate Pathway (PPP)

The main metabolic function of the PPP is to provide anabolic intermediates such as ribulose 5-phosphate, erythrose 4-phosphate, and NADPH for biosynthesis and cell growth. The flux through the nonoxidative PPP in *S. cerevisiae* was found to be much lower than in other yeasts [58], which is sometimes interpreted to be a result of prolonged selection of *S. cerevisiae* for carbon dioxide and ethanol production from

hexose sugars. However, PPP activity is a crucial part of pentose metabolism, since it is virtually the only way to introduce xylulose into the central metabolism.

The insufficient flux through the nonoxidative PPP in *S. cerevisiae* has been confirmed by the superior pentose utilization by strains in which the enzymes of the nonoxidative PPP [ribulose 5-phosphate isomerase (RKI1), ribulose 5-phosphate epimerase (RPE1), transketolase (TKL1) and transaldolase (TAL1)] have been overexpressed. Overexpression of the endogenous *S. cerevisiae* transaldolase (*TAL1*) resulted in improved growth on xylose [49]. It has been shown that the transaldolase from *P. stipitis* is a better gene target, and gene overexpression of *PsTAL1* did not cause growth inhibition when cells were grown on glucose, unlike that of the *S. cerevisiae* *TAL1* [18].

Later, the overexpression of all four nonoxidative PPP genes was shown to improve xylulose consumption by *S. cerevisiae* [59, 60]. Moreover, the improvement resulting from the overexpression of the four genes was higher than when each gene was overexpressed alone [61]. The simultaneous overexpression of the whole nonoxidative PPP, together with *GRE3* deletion, allowed growth on xylose in strains carrying a bacterial XI [62] and in a strain carrying the *Piromyces* XI (strain RWB217, Table 1-2) [20]. PPP overexpression also allowed superior xylose fermentation rates in combination with high levels of XR and XDH (strains TMB3057, Table 1-2) [42, 62].

1.6 *GRE3* Deletion

S. cerevisiae strains reduce xylose to xylitol with an endogenous xylose (aldose) reductase encoded by the *GRE3* gene [63] and other genes putatively encoding enzymes

capable of xylose reduction (*GCY1, YPR1, YDL124W, YJR096W*) [64]. Xylitol strongly inhibits the activity of XI [24, 65], and therefore deleting the *GRE3* gene minimizes xylitol production and improves xylose assimilation in *S. cerevisiae* strains using the XI pathway [66]. Ethanol yields were also noticed to improve for XI-carrying strains with reduced xylitol yields [20] (strain RWB217 v.s. RW 202-AFX in Table 1-2). However, the aldose reductase endoded by *GRE3* belongs to a group of generally stress-induced proteins [63] and the deletion of it reduces the growth by 30% [67], which makes *GRE3* deletion unfavorable in industrial applications.

1.7 Evolutionary engineering

Evolution through adaptation and natural selection is an iterative process of genetic diversification and functional selection. It is nature's design algorithm, and is believed that all natural biological structures were "designed" through this process so that they would best "fit" in certain conditions. Evolutionary engineering is simply the application of this first principle of biological design [68]. The use of evolutionary engineering has proven to be very valuable for obtaining phenotypes of microorganisms with improved properties, such as an expanded substrate range, increased stress tolerance, and efficient substrate utilization [69]. Also, for *S. cerevisiae*, evolutionary engineering has been extensively used to select for industrially relevant phenotypes in addition to rational metabolic engineering, by prolonged cultivation of recombinant *S. cerevisiae* strains in either anaerobic chemostat or sequential batch cultures.

Evolutionary engineering have been successfully employed to obtain improved xylose-utilizing strains (TMB3400[70], C1, C5[71], BH42[62], RWB202-AFX[19],

RWB218[21]), arabinose-utilizing [72, 73] and xylose-arabinose co-utilizing [74, 75] *S. cerevisiae* strains. The best reported xylose-utilizing strain RWB218 was evolved from engineered strain RWB 217(Table 1-2) through two stages: 85 generations in a xylose-limited chemostat for xylose-uptake kinetics, followed by 35 cycles of sequential batch reactor on mixtures of glucose and xylose. A single colony, RWB 218, was then isolated from the sequential batch cultivation and on xylose alone it has the highest reported specific anaerobic growth rate, xylose consumption and ethanol production rate to date (Table 1-2).

Some of the resultant strains have been analyzed in order to identify molecular traits related to the improved ethanolic fermentation of pentose sugars. High-throughput technologies, such as transcription analysis [28, 72, 76], enzyme and metabolite analysis [77] and inverse metabolic engineering [18] have been used. In many cases, the mutations and alterations observed in mutant strains match the rational engineering approach, confirming previous knowledge and hypotheses about control and regulation of pentose metabolism. So far, no novel information has been obtained from high-throughput molecular analyses that would greatly expand our knowledge about pentose metabolism.

1.8 Xylose transporters

S. cerevisiae takes up xylose via its hexose transporters (Hxt1p–Hxt17p, Gal2p, Snf3p), which function by facilitated diffusion [78-80]. Many hexose transporters share a similar structure, consisting of the most conserved 12 putative transmembrane domains, and the most divergent in length and sequence are the cytosolic amino- and carboxyl-

terminal regions [81-83] (Figure 1-2), And some of them show significant sequence homology [84].

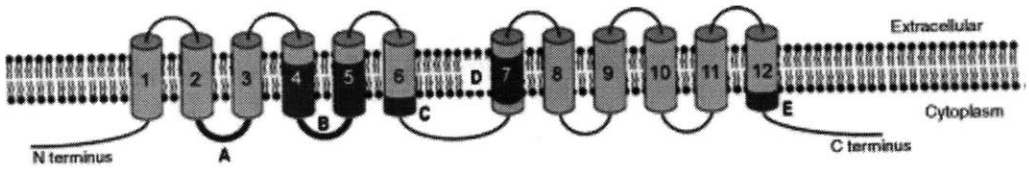


Figure 1-2. Topology of yeast sugar transporters[80]

Engineered *S. cerevisiae* mainly uses the high-affinity system, e.g., *Hxt4, Hxt5, Hxt7 and Gal2*, for xylose transport [85], but the HXTs exhibit a significantly lower affinity for xylose (K_m=49-300 mM) than for glucose (K_m=1-28 mM) [7, 86]. Therefore, glucose strongly inhibits the transport of xylose and restricts xylose assimilation [86, 87].

A metabolic control analysis study demonstrated that transport controlled xylose conversion only at low xylose concentrations [88]. Overexpression of *S. cerevisiae* hexose transporters in a xylose-utilizing yeast strain did not result in faster growth or enhanced xylose fermentation, when the strain was not thoroughly engineered in the xylose metabolism pathway [85]. However, during the evolutionary engineering of RWB 217, all the known enzymes for xylose assimilation overexpressed [21], and the improved xylose uptake characteristics (v_{max} doubled and K_m decreased by 25%) had significantly contributed to xylose fermentation of the resulting strain, RWB 218.

It was predicted that transporters with higher affinity to xylose exist in native xylose-metabolizing yeasts, but few reports exist on expression of pentose transporters in *S. cerevisiae* strains with limited success, although substantial research effort has been made in attempting to isolate fungal xylose transporters [84, 85, 89-94]. This is in part due to the difficulty in actively expressing heterologous membrane proteins. The first

active heterologous expression of a glucose/xylose facilitated diffusion transporter (GXF1, K_m for xylose ~0.4 mM, Table 1-3) and a glucose/xylose symporter (GXS1) from *Candida intermedia* in *S. cerevisiae* was reported [84]. But neither transporter was able to support vigorous growth of the *S. cerevisiae* strain on xylose. However, the expression of Gxf1p in recombinant *S. cerevisiae* was recently shown to improve xylose fermentation [94].

Table 1-3 Biochemical characteristics of pentose transporting proteins [95]

Organism	Transporter	Type[a]	Substrate(s)	K_M (mM)[b]	Reference[c]
Prokaryote					
Escherichia coli	AraE	Symp	ara	$150–320\times10^{-3}$	Daruwalla et al. 1981
	AraFGH	ABC	ara	$4.1–6.1\times10^{-3}$	
	XylE	Symp	xyl	$63–170\times10^{-3}$	Henderson 1990
	XylFGH	ABC	xyl	$0.2–4.0\times10^{-3}$	Sumiya et al. 1995
Bacillus subtilis	AraE		ara, gal, xyl	ND	Krispin and Allmansberger 1998
Eukaryote					
Saccharomyces cerevisiae	Gal2	Unip	ara, gal, glu, xyl	ND (xyl)	Hamacher et al. 2002; Kou et al. 1970
	Hxt1	Unip	fru, glu, man, xyl	880 (xyl)	Saloheimo et al. 2007
	Hxt2	Unip	fru, glu, man, xyl	260 (xyl)	
	Hxt4	Unip	fru, glu, man, xyl	170 (xyl)	
	Hxt5	Unip	fru, glu, man, xyl	ND (xyl)	
	Hxt7	Unip	fru, glu, man, xyl	130 (xyl)	
Pichia stipitis	SUT1	Unip	fru , glu, xyl	1.5 (glu), 145 (xyl), 36 (fru)	Weierstall et al. 1999
	SUT2	Unip	glu, xyl	1.1/55 (glu)[d], 49 (xyl)	
	SUT3	Unip	fru, gal, glu, xyl	0.8/31(glu)[d], 103 (xyl), 49(fru), 176 (gal)	
Candida intermedia	GXS1	Symp	glu, xyl	0.2 (glu), 0.4 (xyl)	Leandro et al. 2006
	GXF1	Unip	glu, xyl	2 (glu), 49 (xyl)	

Heterologous sugar transporters from organisms other than yeast may also function in yeast, if properly expressed and targeted to the plasma membrane. *C. kessleri* Hup1 and *A. thaliana* Stp2 and Stp3 were functionally expressed in the plasma membrane, as they supported growth of the hexose-transport-deficient yeast strain on glucose but not on xylose [96]. None of the *E. coli xylE* and the predicted *A. thaliana* xylose-proton symporters were able to support growth on xylose, most likely because these transporters are not correctly targeted to the plasma membrane in a functional form [85]. Nevertheless,

the expression of heterologous xylose transporters remains a feasible metabolic engineering strategy to further increase the rate of xylose.

1.9 Thesis objective and overview

Efficient utilization of xylose is critical for the cost-effective bioconversion of lignocelluloses. The primary objective of this thesis is to construct *S. cerevisiae* strains that ferment xylose rapidly and thoroughly with a high yield of ethanol, to solve the current limitations in biofuel production from lignocellulosic biomass.

In Chapter 2, xylose metabolism pathways were established in *S. cerevisiae* together with the amplification of the downstream pentose phosphate pathway. The performance of the resulting strains was evaluated in xylose batch cultivation, in order to compare the two pathways and identify any limitation of xylose utilization.

In Chapter 3, evolutionary engineering approaches were used to obtain *S. cerevisiae* strains with promoted xylose-utilization. The multiple-stage evolution included aerobic and oxygen-limited batch cultivations, followed by anaerobic chemostat. Xylose consumption was correlated to the growth rate during the process and quick xylose-utilizing strains were isolated from the evolution, for further evaluation.

In Chapter 4, inverse metabolic engineering approach was designed and employed in the metabolic and evolutionary engineered xylose-utilizing strains, to identify genotypes that limit xylose fermentation. This was done by complementing the unevolved *S. cerevisiae* strain with the genomic library of an evolved strain, and screening on a xylose medium either by growth or a microfluidic-device-based screening system.

In Chapter 5, the engineered xylose-utilizing strains were genomically and biochemically characterized. An in-depth study on the gene duplication of xylose isomerase (XI) was launched to confirm the importance of XI expression in xylose-utilizing strains.

In Chapter 6, the xylose fermentation was optimized on the metabolically and evolutionarily engineered strain. Several other issues were also addressed that could potentially improve the xylose fermentation, such as xylose transporters, other xylose isomerases and a non-phosphorylating GAP dehydrogenase (*gapN*).

Chapter 7 concludes this thesis and discusses the significance of current and future work.

CHAPTER 2. **Construction of Xylose-utilizing Strains**

2.1 Introduction

To enter the central carbon metabolism, xylose must first be converted to xylulose 5-phosphate, an intermediate compound of the pentose phosphate pathway (PPP) (Figure 1-1). Two different pathways can be introduced in *S. cerevisiae* to provide xylose utilization: the xylose reductase/xylitol dehydrogenase (XR-XDH) and xylose isomerase (XI) pathways. The xylulokinase (XK) and the non-oxidative pentose phosphate pathway (PPP) are also necessary for quick pentose assimilation.

In this Chapter, xylose-utilizing strains were successfully constructed, starting from a haploid lab strain BF264-15D. Genes encoding the non-oxidative PPP were over-expressed through integrating an extra copy of each gene into the chromosomes. The two different xylose metabolic pathways, XR-XDH-XK (*XYL1-XYL2-XYL3*) and XI-XK (*XYLA-XYL3*), were introduced via a multiple-copy 2μ plasmid. The resulting strains, H131-XYL123 and H131-XYLA31, were verified through PCR and qRT-PCR in terms of the expression of the key genes for xylose metabolism. However, both strains showed only a slow aerobic growth rate (0.031±0.022 h^{-1} for H131-XYLA31 and 0.081±0.052 $h^{-1}$for H131-XYL123), leaving room for further improvement.

2.2 Materials and Methods

2.2.1 Strains and maintenance

Yeast and bacterial strains (summarized in Table 2-1) were stored in 15% glycerol at -80°C. *E. coli* was grown in Luria-Bertani medium, and 100 mg/L of ampicillin was added to the medium when required.

Table 2-1. *S*trains used in this chapter

Strains	Characteristics	Reference
Escherichia coli DH5α	F⁻ ϕ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *end*A1 *hsd*R17(r_K^-, m_K^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ⁻	Invitrogen 18265-017
Pichia stipitis	Wild type	ATCC 58376
S. cerevisiae		
BF264-15Dau	*MATa, ade1, his2, leu2-3,112, trp1-1, ura3, arg4*	[97]
H110	BF264-15Dau, *TRP1::TDH3p-RKI1-CYC1t-TDH3p-RPE1-CYC1t*	This study
H120	H110, *HIS2::TDH3p-TKL1-CYC1t*	This study
H131	H120, *ADE1::TDH3p-PsTAL1-CYC1t*	This study
H131-XYL123	H131 (pRS426-*XYL123*)	This study
H131-XYLA31	H131 (pRS426-*XYLA3*)	This study

Wild-type yeast strains were routinely cultivated at 30°C in YPD or YPX mediums (10g/L Yeast Extract, 20g/L Peptone and 20g/L D-glucose or D-xylose). SD (Synthetic Defined, Table 2-2) media with appropriate drop-out and glucose (SDG) or xylose (SDX) were used as selective media for cultivation and screening during strain construction. For strain characterization, a minimum Dropout Base (DOB) medium with glucose (YNBG) or xylose (YNBX) and a mixture of appropriate nucleotides and amino acids were used. To select for yeast transformants with the auxotrophic markers or isolate colonies, 20g/L Agar was added to appropriate mediums to make Petri dishes. All yeast media were from Becton, Dickinson and Company, D-glucose (AR-ACS) was from Mallinckrodt Baker, and D-xylose (≥99%, cat#: X1500) was from Sigma-Aldrich.

Table 2-2. Yeast Synthetic defined media

Description	mg/l	Product
Potassium Phosphate Mono	1000.000	YNB
Magnesium Sulfate	500.000	YNB
Sodium Chloride	100.000	YNB
Calcium Chloride	100.000	YNB
Biotin	0.002	YNB
Calcium Pantothenate	0.400	YNB
Folic Acid	0.002	YNB
Inositol	2.000	YNB
Niacin	0.400	YNB
p-Aminobenzoic Acid	0.200	YNB
Pyridoxine Hydrochloride	0.400	YNB
Riboflavin	0.200	YNB
Thiamine HCl	0.400	YNB
Boric Acid	0.500	YNB
Copper Sulfate	0.040	YNB
Potassium Iodide	0.100	YNB
Ferric Chloride	0.200	YNB
Manganese Sulfate	0.400	YNB
Sodium Molybdate	0.200	YNB
Zinc Sulfate	0.400	YNB
TOTAL	**1705.844**	DOB
Ammonium Sulfate	5000.000	DOB
Dextrose	20000.000	DOB
TOTAL	**26705.844**	SD
Adenine	10.000	CSM
L-Arginine HCl	50.000	CSM
L-Aspartic Acid	80.000	CSM
L-Histidine HCl	20.000	CSM
L-Isoleucine	50.000	CSM
L-Leucine	100.000	CSM
L-Lysine HCl	50.000	CSM
L-Methionine	20.000	CSM
L-Phenylalanine	50.000	CSM
L-Threonine	100.000	CSM
L-Tryptophan	50.000	CSM
L-Tyrosine	50.000	CSM
Uracil	20.000	CSM
L-Valine	140.000	CSM
TOTAL	**790.000**	CSM
TOTAL	**27495.844**	SD

Strains from frozen stock were usually streaked out on a plate and single colonies were picked for further cultivation. Colonies picked from Petri dishes, for either strain activation or transformation, were inoculated into 5mL of appropriate medium in a 14 mL BD Falcon™ Round-Bottom Tube using a sterile toothpick. For shake flask cultivation, cells were grown in 50 mL of medium in a 250 mL Erlenmeyer flask shaken at 200 rpm, with sponge plugs used for aerobic growth, and 100 mL of medium and rubber stoppers

(size 6, with 23G1 syringe needle to release CO_2 pressure during fermentation) were used for micro-aerobic or anaerobic growth

Table 2-3. Plasmids used in this chapter

Plasmids	Characteristics	Reference
pRS404	YIp, *TRP1*	[98]
pRS404-*RKI1-RPE1*	*TRP1, TDH3p-RKI1-CYC1t, TDH3p-RPE1-CYC1t*	This study
pRS416GPD	CEN, *URA3, TDH3p*-MCS-*CYC1t*	[99]
pRS416GPD-*RPE1*	pRS416GPD, *Xba* I/*Spe*I - *ScRPE1* - *Xho* I*	This study
pRS416GPD-*RKI1*	pRS416GPD, *Xba* I - *ScRKI1* - *Xho* I	This study
pRS416GPD-*TKL1*	pRS416GPD, *Xba* I - *ScRPE1* - *Xho* I	This study
pRS416GPD-*PsTAL1*	pRS416GPD, *Xba* I/*Spe*I - *PsTAL1* - *Xho* I	This study
pRS416GPD-*XYL1*	pRS416GPD, *Xba* I/*Spe*I - *PsXYL1* - *Xho* I	This study
pRS416GPD-*XYL2*	pRS416GPD, *Xba* I/*Spe*I – *PsXYL2* - *Xho* I	This study
pRS416GPD-*XYL3*	pRS416GPD, *Xba* I/*Spe*I – *PsXYL3* - *Xho* I	This study
pRS426	2μ, URA3	[99]
pRS426GPD	2μ, URA3, *TDH3p*-MCS-*CYC1t*	[99]
pRS426GPD-GPD*XYL3*	pRS426GPD, *Xba* I-*XYLA-Xho* I	This study
pRS426-*XYL1-XYL2*	pRS426, *TDH3p-Spe* I-*PsXYL1-Xho* I-*CYCt-Pst* I - *TDH3p-Xba* I/*Spe*I -*PsXYL2-Hind* III-*CYCt*	This study
pRS426-*XYL123*	pRS426-*XYL1-XYL2, Kpn* I-*TDH3p-PsXYL3-CYC1t-Kpn* I	This study
pRS426-*XYLA31*	pRS426, *TDH3p-XYLA-CYC1t, TDH3p-PsXYL3-CYC1t*	This study
pUC19	Cloning vector, Amp^r	ATCC 37254
pUCAD2	pUC19, *Hind*III-*ADH1-Hind*III	This study
pUCAD2-*PsTAL1*	pUCAD2, *Pst* I-*TDH3p-PsTAL1-CYC1t-Sac* I	This study
pUCHI2	pUC19, *Hind*III-*HIS2-Hind*III	This study
pUCHI2-*TKL1*	pUCHI2, *Pst* I-*TDH3p-TKL1-CYC1t-Kpn* I	This study

*vector was digested with *Xba* I and insert digested with *Spe* I, *Xba* I and *Spe* I are compatible

2.2.2 Enzymes, primers and chemicals

Restriction enzymes, polymerase, DNA-modifying enzymes and other molecular reagents were obtained from New England Biolabs. All general chemicals were purchased from Sigma-Aldrich unless otherwise stated. Synthesized primers for PCR and sequencing were purchased from Invitrogen (Table 2-4). Qiaprep spin miniprep kit was

used for *E .coli* plasmid extraction. Zymoprep™ Yeast Plasmid Miniprep I kit with Zymolyase from Zymo Research were used for yeast DNA preparation. Subcloning efficiency or MAX efficiency DH5α chemically competent *E. coli* were from Invitrogen.

Table 2-4. Primers used in this study

Oligo name	Sequence
RPE1-F-Spe I	AAAAACTAGTATGGTCAAACCAATTATAGC
RPE1-R-Xho I	ATACTCGAGCTAATCTAGCAAATCTCTAGAA
RKI1-F-Xba I	ATTATCTAGAATGGCTGCCGGTGTCCCAA
RKI1-R-Xho I	ATAACTCGAGTCACTTTTCGGTAACTTCA
TKL1-F-Xba I	ATTATCTAGAATGACTCAATTCACTGACAT
TKL1-R Xho I	ATATCTCGAGTTAGAAAGCTTTTTTCAAAGG
PsTAL1-F-Spe I	GGACTAGTATGTCCTCCAACTCCCTTGA
PsTAL1-R-Xho I	TACTCGAGTTAGAATCTGGCTTCCAATTGT
PsXYL1-F-Spe I	CCACTAGTATGCCTTCTATTAAGTTGAACTCTG
PsXYL1-R-Xho I	AACTCGAGCAACCTTCTTAGACGAAGATAG
PsXYL2-F-Spe I	ATACTAGTATGACTGCTAACCCTTCCTTGGTG
PsXYL2-R-Xho I	TTCTCGAGCCAAGCGGTTGACTTACTCA
PsXYL2-R-Hind I	AAAAAAAAGCTTGACTTACTCAGGGCCGTCAATG
PsXYL3-F-Spe I	AAGTTTACTAGTAAAATGACCACTACCCCATTTG
PsXYL3-R-Xho I	TTCTCGAGTTAGTGTTTCAATTCACTTTCCATCTT
XYLA-F-Xba I	AATCTAGAATGGCTAAAGAGTACTTCCCA
XYLA-R-Xho I	TTCTCGAGTTATTGATACATTGCGACAATAG
PsXYL3-F-Spe I	AAGTTTACTAGTAAAATGACCACTACCCCATTTG
PsXYL3-R-Xho I	TTCTCGAGTTAGTGTTTCAATTCACTTTCCATCTT
TDH3p-F-Sac I	GCTGGAGCTCAGTTTATCATTATCA
TDH3p-F-EcoR I	GCTGGAATTCAGTTTATCATTATCAATACTC
TDH3p-F-Pst I	GCTGCTGCAGAGTTTATCATTATCAATACTC
TDH3p-F-Kpn I	GCTGGGTACCAGTTTATCATTATCAATACTC
CYC1t-R-EcoR I	ACTATGAATTCGGCCGCAAATTAAAGCCTT
CYC1t-R-KpnI	TGGGTACCGGCCGCAAATTAAA
CYC1t-R-Pst I	TTGCTGCAGGGCCGCAAATTAAAGCCT

Underline indicates restriction site

2.2.3 Yeast transformation

Yeast strains were transformed with lithium acetate/PEG transformation method.

1) Buffers:
   a) PLI buffer: 1mL LiAc (1M), 1mL H_2O and 8mL PEG 3350 (50%);
   b) ssDNA: salmon sperm (1mg/mL) denatured at 95°C for 5 minutes, and placed immediately on ice;
   c) 0.1M LiAc prepared fresh from 1M LiAc stock.

2) Procedure:
   a) Grow cells overnight in 5mL YP or selective medium, in 14 mL BD Falcon™. Round-Bottom Tube;
   b) Day 2, inoculate in 50 mL in 250 mL Erlenmeyer flask at OD_{600} 0.1-0.2, grow for 2-3 generations or until OD_{600} reached 0.5-1;
   c) Centrifuge culture in a 50mL falcon tube, aspirate medium;
   d) Add 200-400 µL of 0.1M LiAc, depending on the number of cells;
   e) Re-suspend gently and leave at room temperature for 10 minutes, flick tubes 2-3 times;
   f) Prepare solution in an Eppendorf tube in the following order:
      i) 50µL of cells;
      ii) ≤20 µL DNA (≥1 ug, PCR or digestion products don't have to be purified);
      iii) 300 µL of PLI
      iv) 5 µL of ssDNA
   g) Vortex briefly to mix;
   h) Incubate at 42 °C for 25 minutes;

i) Centrifuge at 2000×g for 3 minutes;

j) Aspirate supernatant, re-suspend in sterile PBS and then plate. (If cells need to build resistance to antibiotic, incubate with YPD for at least 5 hours up to overnight, wash with sterile PBS, then plate.)

Usually transformed yeast strains form colonies within 48 hrs on SDG plates; on YNBG plates one more night might be required. On xylose plates (SDX and YNBX) colony forming could take 3-10 days, depending on the xylose-utilizing ability of the strain. If more than 4 days incubation were required, plates were put in another semi-airtight container with an adequate amount of free water (emulate of a moist air incubator) to avoid allowing the agar to dry up. The typical transformation efficiency is 10^2-10^4 CFU/ug DNA transformed, depending on the size and quality of DNA transformed as well as the nature of recombination (plasmid or chromosomal integration, length of homologous sequence, etc.).

2.2.4 Determination of nucleotides, metabolites, cell density and culture dry weight

The concentration and purity of DNA and RNA were determined by using a NanoDrop 1000 Spectrophotometer (Thermo Scientific).

The main metabolites D-glucose, D-xylose, xylitol, ethanol, glycerol and acetic acid were detected by a Waters Alliance 2695 HPLC with a Waters 410 Differential Refractometer and a Bio-Rad HPX-87H column. The column was eluted at 50°C with 14 mM of sulfuric acid at a flow rate of 0.7 mL/min.

Cell density was determined by measuring absorbance at 600 nm (OD_{600}) against a medium blank in a 10 mm cell using an Ultrospec 2100 pro UV/Visible

Spectrophotometer (GE). Specific growth rate (μ_{max}) was calculated using cell growth data points from the early exponential growth phase:

$$\mu_{max} = \frac{d \ln(OD_{600})}{dt} \qquad \textbf{Equation 2-1}$$

For dry cell weight (DCW), culture samples (1-10 mL) were filtered with pre-weighed cellulose acetate syringe filters (diameter 25mm, pore size 0.2 μm; VWR). After filtration of the broth, the cells were washed twice with de-ionized water, dried in a microwave oven for 15 minutes at 500 W, desiccated in a 60°C oven for another 15 minutes, and then weighed. Duplicate determinations varied by less than 2%. DCW was further correlated to OD_{600} for typical strains.

2.2.5 Quantitative PCR for determining gene copy number and expression

Genomic DNA was isolated using Promega Wizard Genomic DNA Purification Kit. Quantitative PCR was performed by using a Bio-Rad iCycler iQ Real-Time PCR Detection System and iQ SYBR Green Supermix. Quantitative PCR conditions were as recommended by the supplier: 400 nM of each primer was used in one step of 95°C for 3 min, and 40 cycles of 95°C for 10 s, 55°C for 15 s and 72°C for 15 s. Endogenous *PGK1* and *ACT1* were used as reference to quantify the copy number of genes of interest; 12.8, 3.2, 0.8, 0.2, and 0.05 ng of DNA template were used.

Quantitative RT-PCR was carried out similarly using the iScript One-Step RT-PCR Kit with SYBR Green (Bio-Rad), under conditions recommended by the supplier. *PGK1* and *ACT1* were also used as reference and 10, 1, 0.1 and 0.01 ng of total RNA template were used. All reactions were performed in triplicate.

2.3 Results and discussion

2.3.1 Measuring of nucleic acids, metabolites and OD_{600} and DCW

1) Nucleic acid analysis

At a wavelength of 260 nm, the extinction coefficient of 0.020 $(\mu g/mL)^{-1} \cdot cm^{-1}$ was used for double-stranded DNA and 0.027 $(\mu g/mL)^{-1} \cdot cm^{-1}$ used for single-stranded DNA and RNA[100]. Thus, an OD_{260} of 1 corresponds to a concentration of 50 μg/mL for dsDNA and 37 μg/mL for ssDNA or RNA. This method of calculation is valid for an OD of up to 2. For a pure DNA sample, the OD_{260}/OD_{280} should be ~2 and OD_{260}/OD_{230} ~1.8. Lower OD_{260}/OD_{280} is usually caused by protein contamination and lower OD_{260}/OD_{230} is common for gel extracted DNA.

2) HPLC chromatogram and standard curves

Usually HPLC assay is able to return reproducible data with <3% error at 95% confidence. With the refractive index detector, the limits of detection (LOD) are ~0.5 g/L for ethanol and acetate, while for glucose, xylose, xylitol and glycerol the LOD are ~0.2 g/L. Table 2-5 and Figure 2-1 show typical chromatogram and retention times for the major metabolites of interest.

Table 2-5. HPLC peaks annotation and HPLC standard curves coefficients

Metabolites	Retention time (min)	a	b	R^2
Solvent	5.83	/	/	/
D-glucose	8.08	1.07522E-05	-0.018	0.99999
D-xylose	8.62	1.0648E-05	-0.131	0.99991
Xylitol	9.82	1.0583E-05	0.009	0.99998
Glycerol	11.77	9.60647E-06	-0.027	0.99989
Acetate	13.19	2.6206E-05	0.162	0.99692
Ethanol	18.99	2.61686E-05	0.104	0.99992

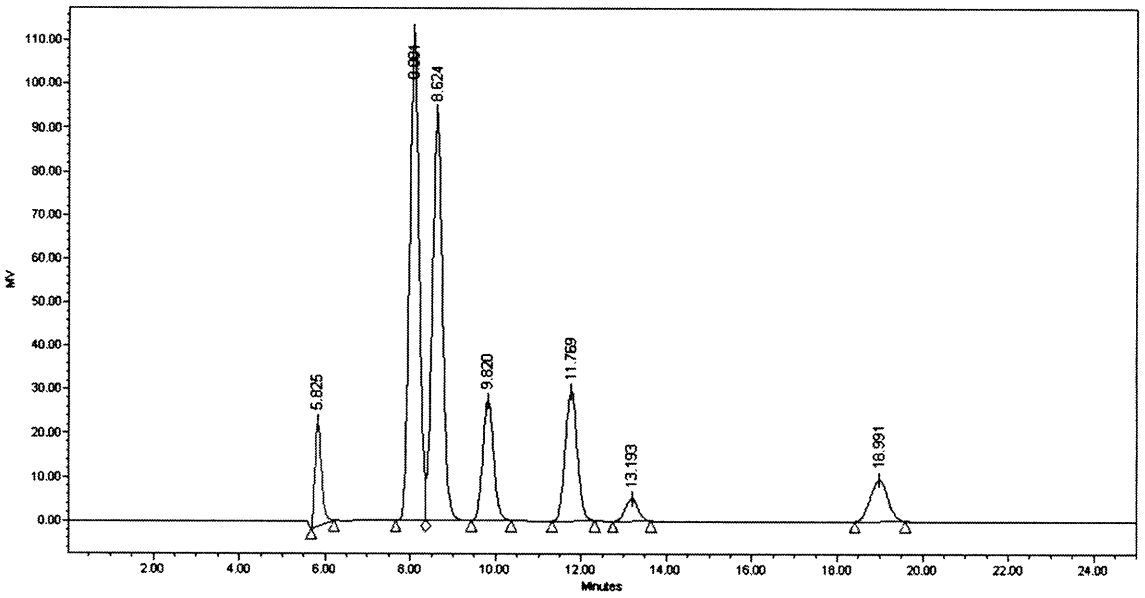


Figure 2-1. HPLC chromatogram

Figure 2-2 and Table 2-5 are the standard curves and coefficients of each metabolite of interest, based on linear regression.

$$[\text{Metabolite}](\text{g/L}) = a \times \text{peaksize} + b \qquad \textbf{Equation 2-2}$$

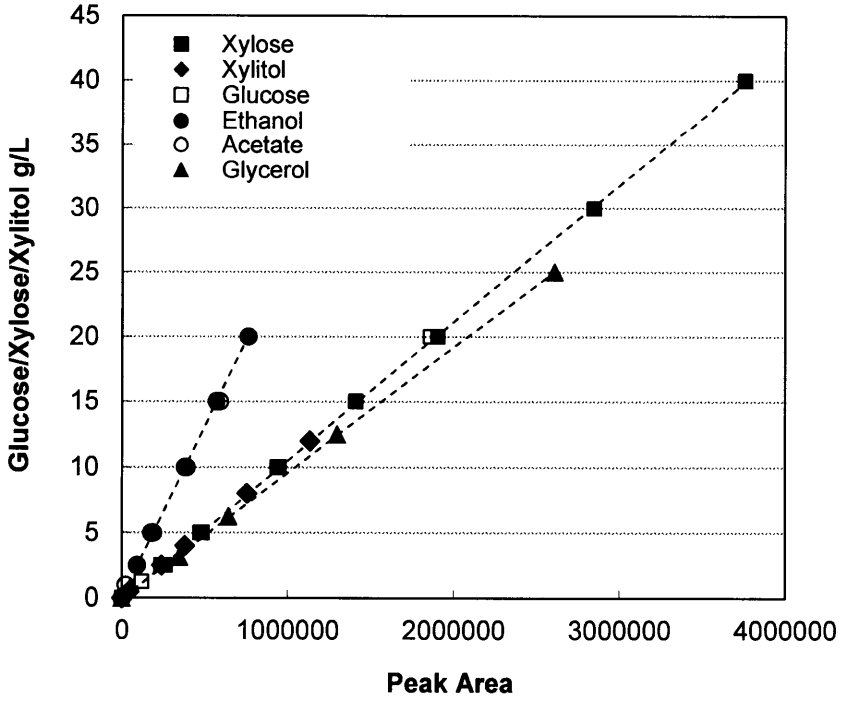


Figure 2-2. HPLC standard curves

The coefficients of the standard curves are usually constant in one batch of samples if the column is well conditioned. They were checked when the new mobile phase solution was made or after a longer period of sampling (over a week), and re-done if necessary.

3) Dry cell weight and OD_{600} correlation

The absorbance at 600 nm of the samples measured in a spectrophotometer is correlated to the dry cell weight. Samples were diluted to appropriate concentrations with absorbance between 0.1 and 0.5 for the highest accuracy. The optical density is also a function of cell morphology such as size and shape, which are dependent on the nutrient composition and the cell growth phase. Consequently, an independent calibration curve is required for each condition in accurate research work.

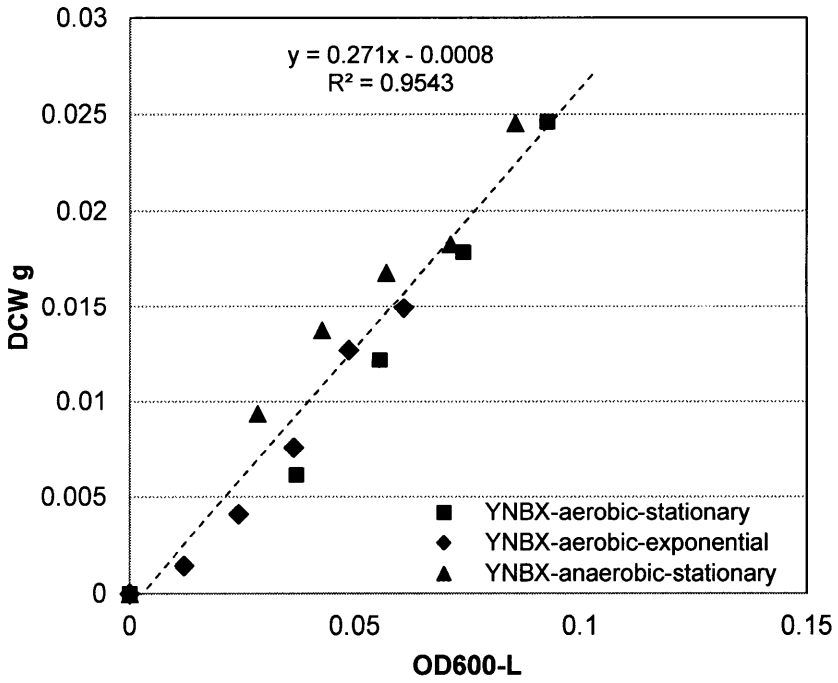


Figure 2-3. Dry cell weight calibration curve

As shown in Figure 2-3, an optical density of 1 unit corresponds to approximately 0.2-0.3 g/L of dry cell, depending on the medium and growth phase. In this research a coefficient of 0.267 g DCW/L·OD_{600} is used for the calculation.

2.3.2 Pentose phosphate pathway over-expression strain construction

Plasmids used in this study are summarized in Table 2-3. All *S. cerevisiae* strains in this study were constructed from the BF264-15Dau parent strain (kindly provided by Gerald R. Fink, Whitehead Institute). In order to integrate the genes of the non-oxidative pentose phosphate pathway, several plasmids were constructed. To put the endogenous *RPE1, RKI1, TKL1* and *Pichia stipitis TAL1* under the strong constitutive glyceraldehydes-3-phosphate dehydrogenase promoter (*TDH3p*), the genes were amplified from *S. cerevisiae* or *P. stipitis* genomic DNA using appropriate primers (Table 2-4), digested by *Xba* I-*Xho* I or *Spe* I-*Xho* I, and ligated into *Xba* I-*Xho* I sites of pRS416GPD, resulting in pRS416GPD-*RPE1*, pRS416GPD-*RKI1*, pRS416GPD-*TKL1* and pRS416GPD-*PsTAL1* (Table 2-3).

4) Integration of *RPE1* and *RKI1*

Expression cassettes *TDH3p-RKI1-CYC1t* and *TDH3p-RPE1-CYC1t* were then PCR amplified from the above plasmids using appropriate primers, digested accordingly and inserted to the *Sac* I-*EcoR* I and *EcoR* I-*Kpn* I sites of pRS404 to yield the integration plasmid pRS404-*RKI1-RPE1*.

The plasmid was linearized by *Hind* III in *TRP1* and transformed to BF264-15Dau. After selection on tryptophan drop-out medium, H110 was achieved.

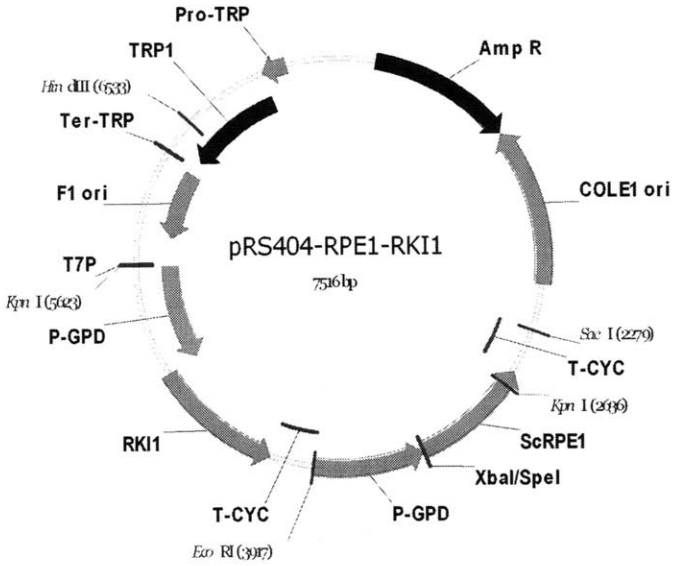


Figure 2-4. Integration plasmid pRS404-*RPE1-RKI1*

5) Integration of *TKL1*

pUCHI2-*TKL1* (Figure 2-5) was constructed in a similar way, *TDH3p-TKL1-CYC1t* was amplified by PCR from pRS416GPD-*TKL1* using primers TDH3p-F-Pst I and CYC1t-R-Kpn I, digested and ligated into *Pst* I-*Kpn* I sites of pUCHI2. The plasmid was linearized by *Hpa* I in *HIS2* auxotrophic marker and transformed to H110 to yield H120 after selection on a histidine drop-out medium.

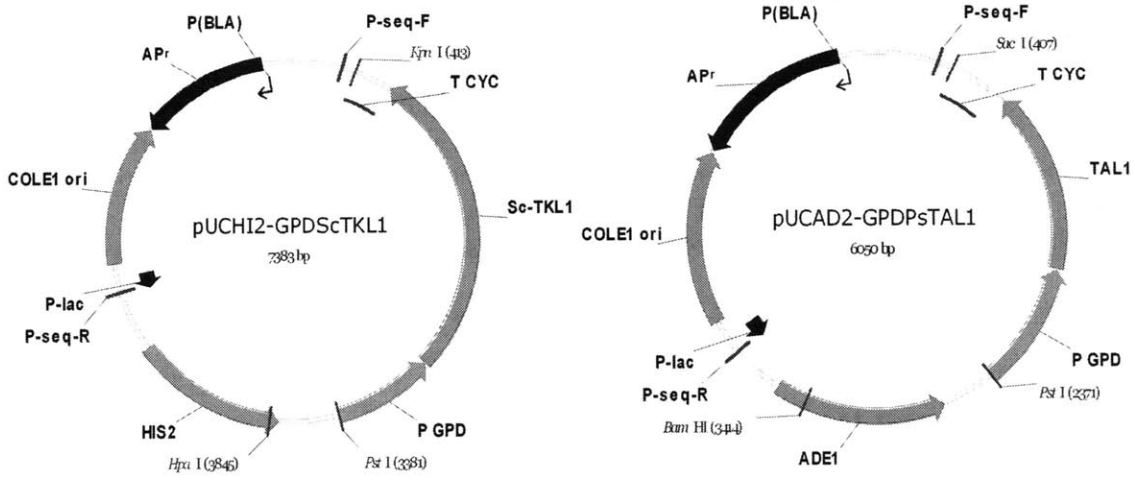


Figure 2-5. Integration plasmids pUCHI2-*TKL1* and pUCAD2-Ps*TAL1*

6) Integration of *PsTAL1*

For the transaldolase (*TAL1*), it has been shown that the *TAL1* from *P. stipitis* is a better gene target than *S. cerevisiae TAL1*. It has higher specific activity and over-expression of *PsTAL1* did not cause growth inhibition when cells were grown on glucose, unlike that of the *S. cerevisiae TAL1* [18].

pUCAD2-*PsTAL1* (Figure 2-5) containing *TDH3p-PsTAL1-CYC1t* at *Sac* I- *Pst* I sites was similarly constructed using pRS416GPD-*PsTAL1* as template, linearized with *BamH* I, and then integrated into the genome of H120 to form the PPP over-expression background strain H131 (*MATa, TRP1::TDH3p-RKI1-CYC1t-TDH3p-RPE1-CYC1t,HIS2::TDH3p-TKL1-CYC1t, ADE1::TDH3p-PsTAL1-CYC1t, leu2, ura3, arg4*).

2.3.3 Xylose metabolic pathway based on XI

For the xylose metabolic pathway based on xylose isomerase (XI), both genes *XYLA* coding for XI and *XYL3* coding for xylulokinase were expressed on a multicopy plasmid pRS426GPD. In order to increase the activity of xylulokinase, the *XYL3* gene from *Pichia stipitis* was expressed in addition to the endogenous *XKS1*, since it was reported to have higher activity [17]. Moreover, *PsXYL3* was also reported to have higher specific activity on xylulose than xylitol, avoiding any possible toxic xylitol-5P production [57]. The *PsXYL3* was amplified from the *P. stipitis* (ATCC 58376) genome using PCR and inserted into pRS416GPD behind the *TDH3* promoter. The *TDH3p-PsXYL3-CYC1t* expression construct was then cloned into pRS426GPD at the *Kpn* I site, resulting in pRS426GPD-GPD*XYL3* (Figure 2-6).

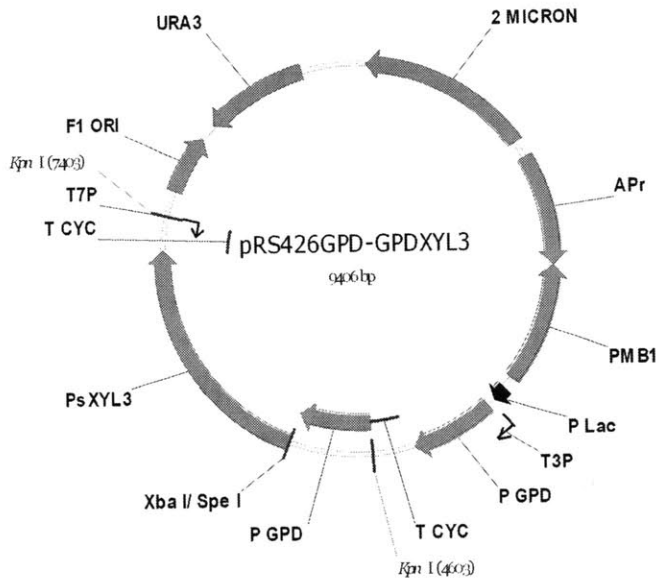


Figure 2-6. Cloning of *PsXYL3*

For the single site insertion, the *TDH3p-PsXYL3-CYC1t* expression construct could have two orientations but only the sense structure (same orientation for the two *THD3* promoters) was achieved.

The gene *XYLA* for XI from *Piromyces* sp. E2 was used for the construction of a xylose metabolic pathway as it was the only reported *XYLA* that could be functionally expressed in *S. cerevisiae*. Since the original source (*Piromyces* sp. E2 ATCC 76762) of the *XYLA* gene had been discontinued, the *XYLA* was codon optimized and synthesized (see Appendix), and then cloned into plasmid pRS426GPD-GPD*XYL3* at the *Spe* I/ *Hind* III sites to yield pRS426GPD*XYLA*-GPD*XYL3* (Figure 2-7, called pRS426-*XYLA3* for short in further discussion).

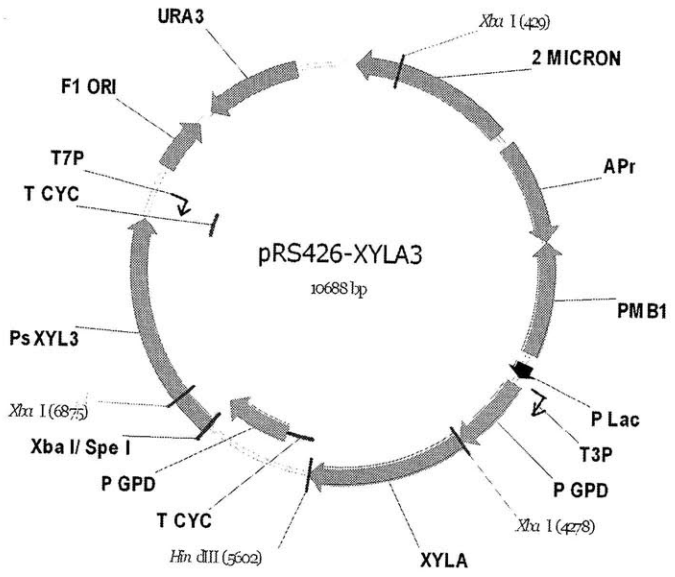


Figure 2-7. Plasmid pRS426-*XYLA3*

2.3.4 Xylose metabolic pathway based on XR-XDH

To construct the fungal redox xylose metabolic pathway, *XYL1* coding for XR and *XYL2* coding for XDH from *P. stipitis* (ATCC 58376) were expressed under strong constitutive *TDH3* promoter, and put on the multiple copy plasmid pRS426 together with *PsXYL3*. *XYL1* and *XYL2* genes were amplified from the *P. stipitis* (ATCC 58376) genome using PCR and inserted into pRS416GPD behind the *TDH3* promoter. *XYL1-CYC1t* and *TDH3p-XYL2* were then PCR amplified from the plasmids and ligated by *Pst* I (Figure 2-8). The ligation product was further amplified using PsXYL1-F-Spe I and PsXYL2-R-Hind III primers, digested accordingly and ligated into plasmid pRS426GPD-GPD*XYL3* constructed above at the *Xba* I/ *Hind* III sites to yield pRS426GPD*XYL1*-GPD*XYL2*- GPD*XYL3* (Figure 2-8, named pRS426*XYL123* for short).

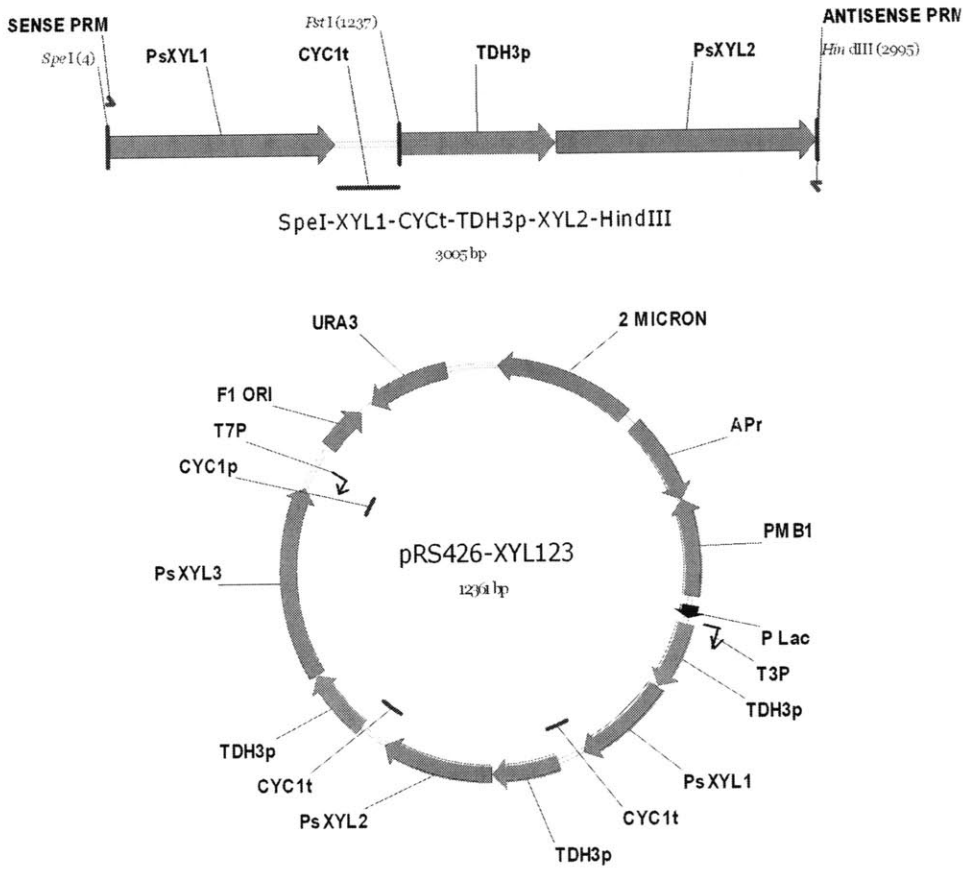


Figure 2-8. Cloning of X*YL1-XYL2-XYL3*

2.3.5 Xylose-utilizing strains construction and characterization

H131 was then transformed with pRS426-*XYLA3* and pRS426-*XYL123*, separately, to achieve xylose-utilizing strains H131-XYLA31 and H131-XYL123. After transformation, single colonies were isolated from selective SDG plate with uracil drop-out. The specific growth rates of H131-XYLA31 and H131-XYL123 in SDX liquid medium were 0.031±0.022 h^{-1} and 0.081±0.052 h^{-1} under aerobic condition. As a control, H131 showed no growth on xylose under similar conditions within an extended period.

After transformation, the xylose reductase/xylitol dehydrogenase (*XYL1-XYL2*) pathway resulted in a higher growth rate on pentose sugar than the combination of the xylose isomerase (*XYLA*) pathway. Similar results have been observed in the literature

when comparing the two xylose metabolic pathways [101]. Heterologous expression of xylose isomerase in *S. cerevisiae* has been proved to be very difficult and was successful only when combined with extensive adaptation protocols [19, 45]. Therefore, further evolution of the xylose-utilizing strains was necessary to improve their performance.

2.3.6 Verification of genes integration and transformation

The successful cloning was verified through PCR (data not shown). For heterologous genes (*XYLA, PsXYL1, PsXYL2, PsXYL3* and *PsTAL1*) the existence of the genes in the engineered strain were confirmed, while for endogenous *TKL1, RPE1* and *RKI1* primer pairs spanning gene ORF and chromosome context were used to verify the proper integration. The transcription of overexpressed genes was checked through qRT-PCR using the endogenous housekeeping gene *ACT1* as reference, as shown in Figure 2-9.

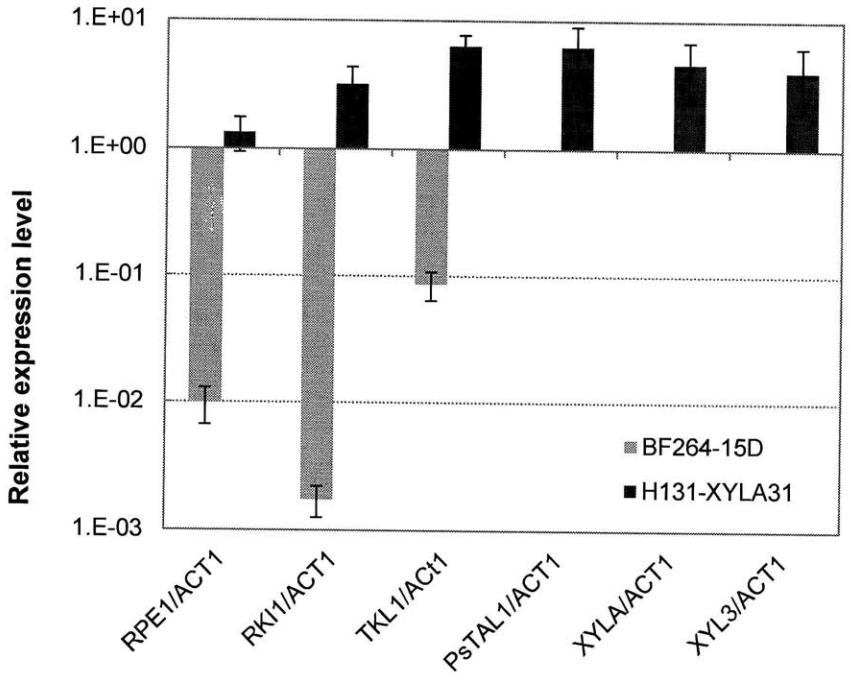


Figure 2-9. Expression of transformed genes

The expression level of *RPE1, RKI1* and *TKL1* were rather low for the background strain BF264-15D growing micro-aerobically on glucose, while for the constructed H131-XYLA31 transcription level of PPP was elevated 1 to 3 orders of magnitude after overexpression. The heterogeneous *PsXYLA, PsXYL3* and *PsTAL1* were confirmed highly expressed in the H131-XYLA31 while showing no expression in the wild-type BF264-15D. *PsXYL1* and *PsXYL2* expression were not assessed by qRT-PCR since their expression could be indirectly verified by the growth of strain H131-XYL123 on xylose medium. Moreover, the strain was not extensively developed due to the difficulty in resolving the *XYL1-XYL2* cofactor imbalance.

2.4 Conclusions

Xylose-utilizing strains were successfully constructed through heterologous expression of either XI-XK (*XYLA-XYL3*) or XR-XDH-XK (*XYL1-XYL2-XYL3*) xylose metabolic pathways, which were named H131-XYLA31 and H131-XYL123 respectively. Pentose phosphate pathway was also enforced to facilitate a high flux of xylose assimilation, by overexpression of the non-oxidative PPP genes. The key genes for xylose metabolism and their transcription were verified via PCR and qRT-PCR. Both strains showed only a slow aerobic growth rate (0.031 ± 0.022 h^{-1} for H131-XYLA31 and 0.081 ± 0.052 $h^{-1}$for H131-XYL123) after transformation, suggesting other expression or regulatory limit in xylose utilization.

CHAPTER 3. Evolutionary Engineering of Xylose-utilizing Strains

3.1 Introduction

For the yeast *Saccharomyces cerevisiae*, evolutionary engineering has been extensively used to select for industrially relevant phenotypes such as an expanded substrate range and increased stress tolerance. The approach was also applied for achieving pentose-utilizing strains [21, 71, 74, 75], especially on those using a XI or a bacterial arabinose pathway. In previously constructed xylose-utilizing strains, the xylose metabolic pathways together with the over-expression of pentose phosphate pathway did not result in efficient xylose consumption, implying that there were other expression and/or regulatory limitations in the system. In this chapter, evolutionary engineering was employed to obtain improved xylose-utilizing strains, in multiple stages.

Started with H131-XYLA31 (Table 3-1, *XYLA-XYL3* pathway on multiple-copy plasmid, one extra copy of non-oxidative PPP genes under strong promoters integrated onto chromosomes), the evolution underwent three stages including aerobic sequential batch cultivation, oxygen-limited sequential batch cultivation and anaerobic chemostat cultivation. Strains were isolated from the evolved populations periodically for evaluation in terms of growth and fermentation. The finally evolved strain, H131E8-XYLA31, displayed significantly improved anaerobic growth rate (0.120 ± 0.004 h^{-1}) and xylose consumption rate (0.916 $g \cdot g^{-1} h^{-1}$) in minimal medium. The performances of H131E8-XYLA31 were comparable to those of the best reported xylose-utilizing strain RWB 218, before any optimization of fermentation.

3.2 Materials and Methods

3.2.1 Shake flask batch cultivation

For the growth rate measurement, cells were grown in 50 mL of SDX or YNBX medium in a 250 -mL Erlenmeyer flask shaken at 200 rpm. Sponge plugs were used for aerobic growth. For micro-aerobic growth, rubber stoppers (with 23G1 syringe needles to release CO_2 pressure during fermentation) were used to block oxygen diffusion, and 100 mL of medium was used to reduce the ratio of the headspace to the aqueous phase. Beyond these approaches, nitrogen flow at about 1 L/min was used to sparge the medium during the inoculation and sampling for about 30 seconds, to achieve an anaerobic condition. Initial cell growth (OD_{600}) was used for calculation of the maximum specific growth rate.

Sequential batch cultivation was carried out under similar aerobic or oxygen-limited conditions in YNBX or SDX medium. Once the culture reached the stationary phase., an inoculum was taken to start a new batch at OD_{600} of 0.05-0.1.

3.2.2 Chemostat cultivation

Anaerobic chemostat cultivation was carried out at 30°C in 500-mL Water Jacketed Spinner Flasks (Bellco Glass) with a fixed working volume of 200 mL, using a leveled effluent driven by a peristaltic pump. Cultures were grown in YNBX medium with various xylose concentrations. After inoculation and completion of the batch phase, chemostat cultivation was initiated by feeding YNBX to the fermentor at increasing dilution rates. The initial dilution rate (0.02 h^{-1}) was increased stepwise every time the culture reached a steady state. Cultures were stirred at 200 rpm. The pH was monitored

but not maintained. The effluent broth was checked periodically by microscopic analysis to verify the absence of contamination, and analyzed to determine biomass and metabolite concentrations.

3.2.3 Anaerobic fermentation of xylose

For the characterization of the engineered and evolved strains, pre-cultures were grown in YNBX. The first pre-culture (5mL in a 14-mL BD Falcon Round-Bottom Tube) grown overnight aerobically was used to inoculate a second 50-mL inoculum in a plugged 250-mL flask. Cells from the late exponential phase were harvested and inoculated at an OD_{600} of about 0.2. Fermentation was conducted in 1.3-liter BioFlo 110 bioreactors with 1 liter of 2×YNBX with 40 g/L xylose. Temperature was set at 30 °C, agitation was set at 200 rpm, and pH was maintained at 5.0 by automatic addition of 2N NaOH. Anaerobic conditions were obtained by sparging ultrapure nitrogen at 1 L/min until DO dropped to zero at the beginning of the fermentation.

3.2.4 Strains, other materials and methods

Yeast strains constructed and evolved in this chapter are summarized in Table 3-1.

Table 3-1. *S. cerevisiae* strains used in this chapter

S. cerevisiae Strains	Characteristics
H131	BF264-15Dau , *MATa, leu2-3,112, ura3, arg4, TRP1::RKI1- RPE1, HIS2::TKL1, ADE1::PsTAL1*
H131-XYL123	H131 (pRS426-*XYL123*), *PsXYL1, PsXYL2, PsXYL3*
H131-XYLA31	H131 (pRS426-*XYLA3*), *Piromyces XYLA, PsXYL3*
H131E1-XYL123	H131-XYL123, isolated from aerobic sequential batch cultivation
H131E1-XYLA31	H131-XYLA31, isolated from aerobic sequential batch cultivation
H131E3-XYLA31	H131E1-XYLA31, isolated from micro-aerobic sequential batch cultivation
H131E5-XYLA31	H131E3-XYLA31, isolated from anaerobic sequential batch cultivation
H131E8-XYLA31	H131E5-XYLA31, isolated from anaerobic chemostat

For other materials and methods such as the analysis of metabolites and cell density please refer to Chapter 2, Materials and Methods.

3.3 Results and Discussion

The engineered *S. cerevisiae* strains H131-XYLA31 and H131-XYL123 (Table 3-1) were used to initiate the evolution. Both strains showed slow aerobic growth on the xylose medium, and therefore make it possible to select for faster growth on xylose. When xylose is used as the solo carbon source, the specific growth rate should be coupled with xylose consumption. Specifically, under oxygen-limited conditions, the growth rate should strictly correlate with both xylose utilization and ethanol production. Moreover, xylose as a non-preferred carbon source for *S. cerevisiae* may serve as a nutritional stress and induce spontaneous mutagenesis and genomic rearrangements, which could further accelerate the evolution process [102-104]. Given these hypotheses, the evolution was deployed in three stages, as described in Figure 3-1 below.

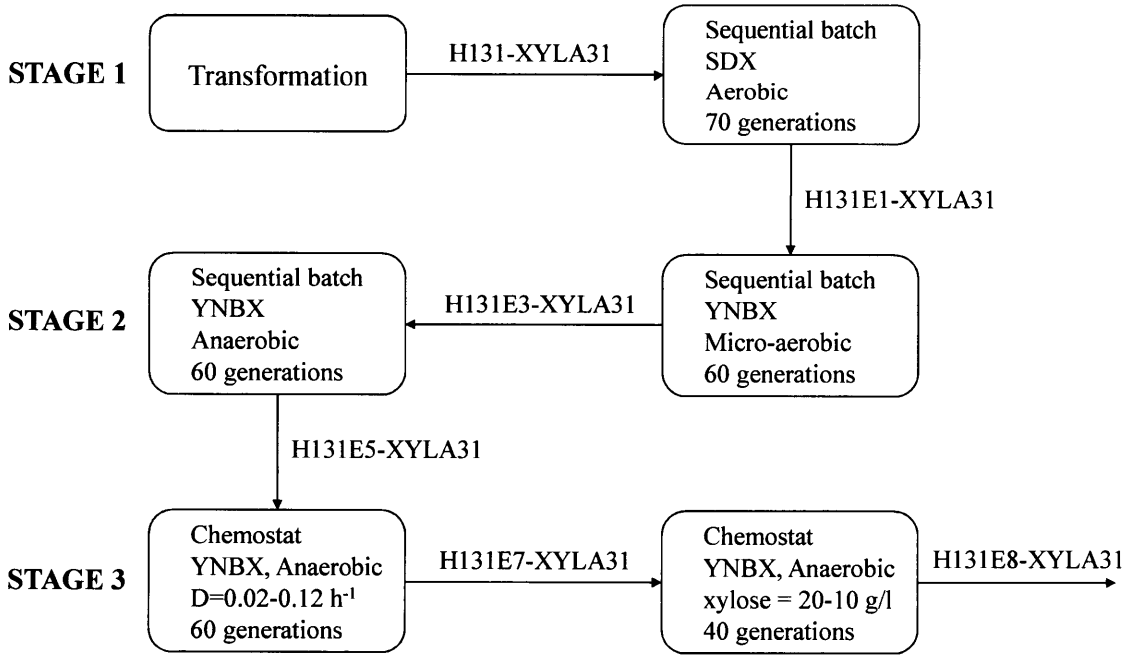


Figure 3-1. Scheme of strain evolution

3.3.1 Aerobic sequential batch cultivation for improved growth on xylose

The first stage of evolution was carried out through aerobic sequential batch cultivation in the relatively rich SDX medium, for both H131-XYL123 and H131-XYLA31. Once the batch cultivation reached the stationary phase, it was re-inoculated into a new batch at OD_{600} of 0.05-0.1. The OD_{600} of the stationary phase cultivation was typically 8~10, and therefore, each batch is approximately equivalent to 7 generations. The evolution lasted about 70 generations, during which the specific growth rates were monitored, as shown in Figure 3-2.

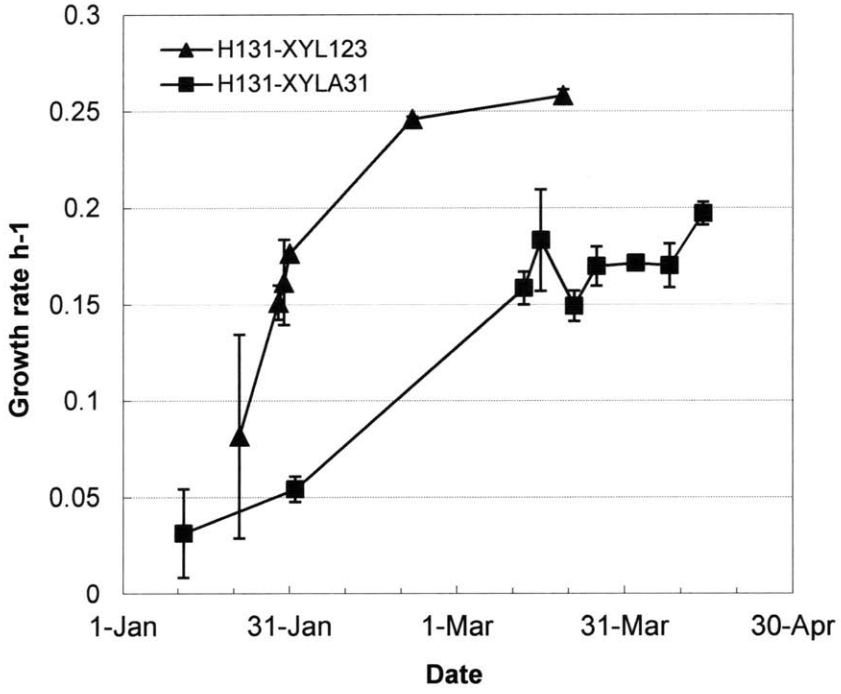


Figure 3-2. Evolution process of strains under aerobic condition

The H131-XYL123 harboring the heterogeneous fungal xylose metabolic pathway adapted relatively more quickly. In about 30-40 generations, a specific growth rate of about 0.25 h^{-1} was achieved on xylose (as a reference, the growth rate on glucose

and similar conditions was 0.4-0.45 h^{-1}). However, the H131-XYLA31 evolution was much slower, and within about 70 generations, a μ_{max} of about 0.2 h^{-1} was achieved.

At the end of this stage, cells from the shake flask cultivations were streaked on an SDX plate for isolation of single colonies according to their sizes. The selected strains H131E1-XYL123 and H131E1-XYLA31 (Table 3-1) were characterized in shake flasks in terms of micro-aerobic fermentation performance, as shown in Figure 3-3 and Figure 3-4.

Fermentations were started with SDX (2% xylose) medium. Within about 80 hours, most xylose was consumed. An extended fermentation was required to deplete xylose due to slow assimilation of xylose under low xylose concentration, when the xylose transport likely limits its uptake. The important characteristics of both strains with respect to micro-aerobic fermentation were calculated and summarized in Table 3-2.

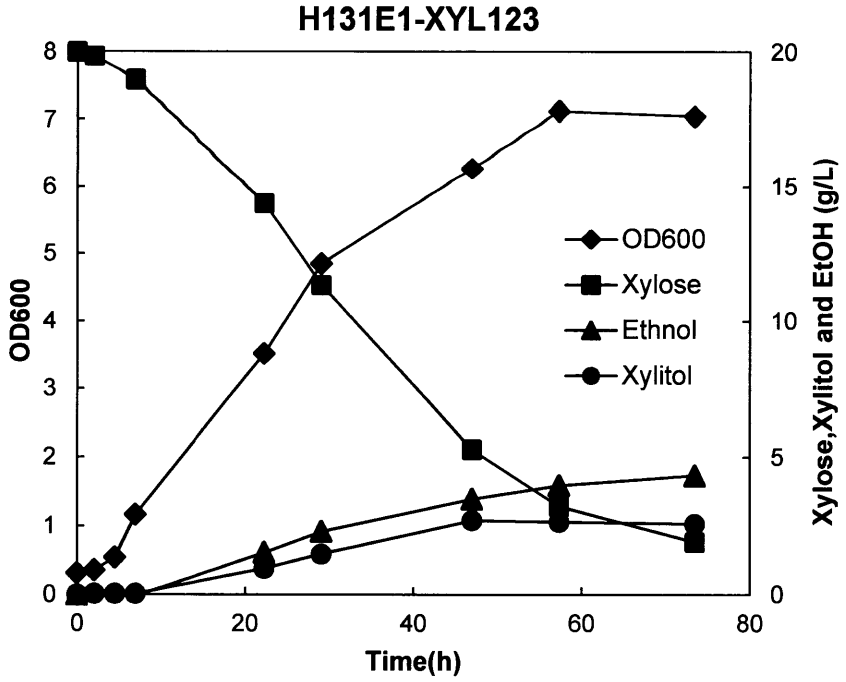


Figure 3-3. Fermentation profiles of H131E1-XYL123

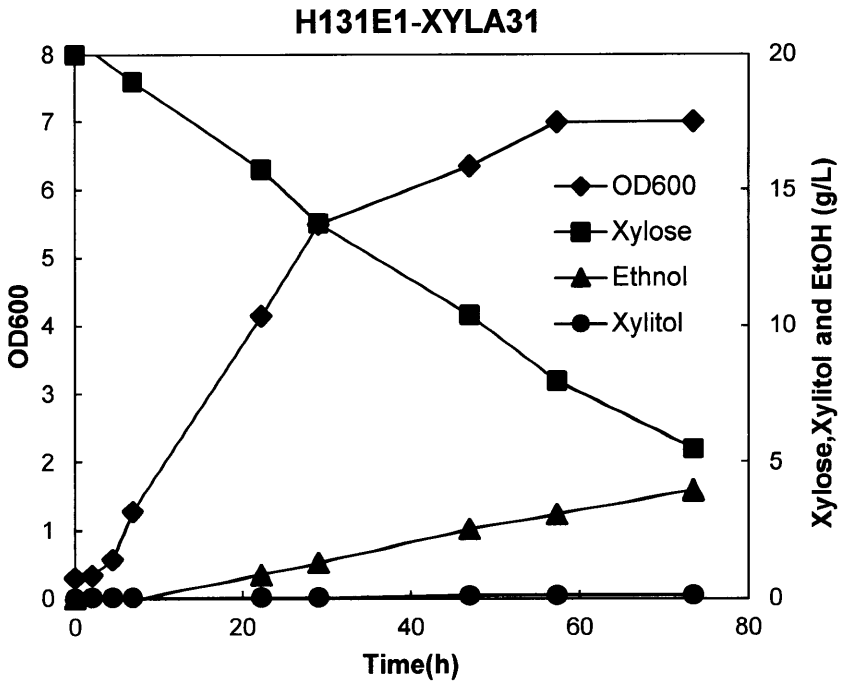


Figure 3-4. Fermentation profiles of H131E1-XYLA31

Table 3-2. Characteristics of H131E1-XYL123 and H131E1-XYLA31 fermentation

		H131E-XYL123	H131E-XYLA31
μ_{max}	h^{-1}	0.245±0.039	0.197±0.006
Δt	h	73.45	73.45
Δ[Xylose]	g/L	18.08	14.51
Δ[Ethanol]	g/L	4.35	3.96
Δ[DCW]	g/L	1.796	1.785
Δ[Xylitol]	g/L	2.582	0.133
$OD_{average}$		4.791	4.988
$DCW_{average}$	g/L	1.28	1.33
q_{xylose}	$g \cdot L^{-1} h^{-1}$	0.246	0.197
	$g \cdot g^{-1} h^{-1}$	0.193	0.148
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	0.059	0.054
	$g \cdot g^{-1} h^{-1}$	0.046	0.041
EtOH yield	g/g	0.241	0.273
Biomass yield	g/g	0.099	0.123
Xylitol yield	g/g	0.143	0.009

Under micro-aerobic conditions, the initial growth resembles aerobic growth since oxygen exists in the headspace of the flask and the biomass concentration is low.

H131E1-XYLA31 showed an aerobic growth rate of 0.197±0.006 h^{-1}, about six-fold higher than the parent strain H131-A31. H131E1-XYL123 exhibited a slightly higher μ_{max} of 0.245±0.006 h^{-1} and a growth profile similar to that of H131E1-XYLA31.

The specific xylose consumption rate of H131E1-XYL123 is about 30% higher than that of H131E1-XYLA31, implying that the fungal *XYL1-XYL2* pathway had a higher efficiency in xylose conversion. This finding is consistent with the quicker adaptation of the H131-XYL123 strain. However, the ethanol productivity of H131E1-XYL123 was only about 12% higher than H131E1-XYLA31, mainly due to the lower ethanol yield and higher xylitol yield of the *XYL1-XYL2* pathway.

Xylitol accumulation is an obvious disadvantage of the *XYL1-XYL2* pathway, particularly under oxygen-limited conditions, owing to the cofactor imbalance of the redox pathway. The xylitol production and yield of H131E1-XYL123 were 2.582 g/L and 0.143 g xylitol/g xylose, respectively, while the xylitol accumulation of the XI-based strain--H131E1-XYL123--was only 0.13 g/L, around the detection limit. The cofactor rebalance is involved in either the protein engineering of XR/XDH (*XYL1/XYL2*) or other complex metabolic network rearrangements, and both approaches had been under development with limited success. On the other hand, the XI (*XYLA*) pathway gets around the cofactor imbalance and is shown to be successfully established in *S. cerevisiae*. Therefore, further research on xylose utilization focused on the XI pathway.

3.3.2 Evolution by oxygen-limited sequential batch cultivation

The evolved population from which the H131E1-XYLA31 was isolated was further evolved micro-aerobically and then anaerobically. A similar sequential batch cultivation setup was applied except that minimal YNBX medium was used instead of the

SDX medium. Each batch was equivalent to about 6 generations because the OD_{600} of the culture in the stationary phase was only 4~5, about half as much as that on the SDX medium. Strains were isolated from the culture periodically to test the performance. After about 60 generations of micro-aerobic selection, a portion of the cells from the batch was streaked on a SDX plate for isolation of single colonies. The strain with best anaerobic specific growth rate was isolated (H131E3-XYLA31, Table 4-2). The sequential batch cultivation was switched to anaerobic conditions for another 60 generations. One strain (H131E5-XYLA31, Table 4-2) was isolated using a similar approach. The selected strains H131E3-XYLA31 and H131E5-XYLA31 were characterized in batch cultivation with respect to anaerobic growth and fermentation on xylose, as shown in Figure 3-5/Table 3-3 and Figure 3-6/Table 3-4, respectively.

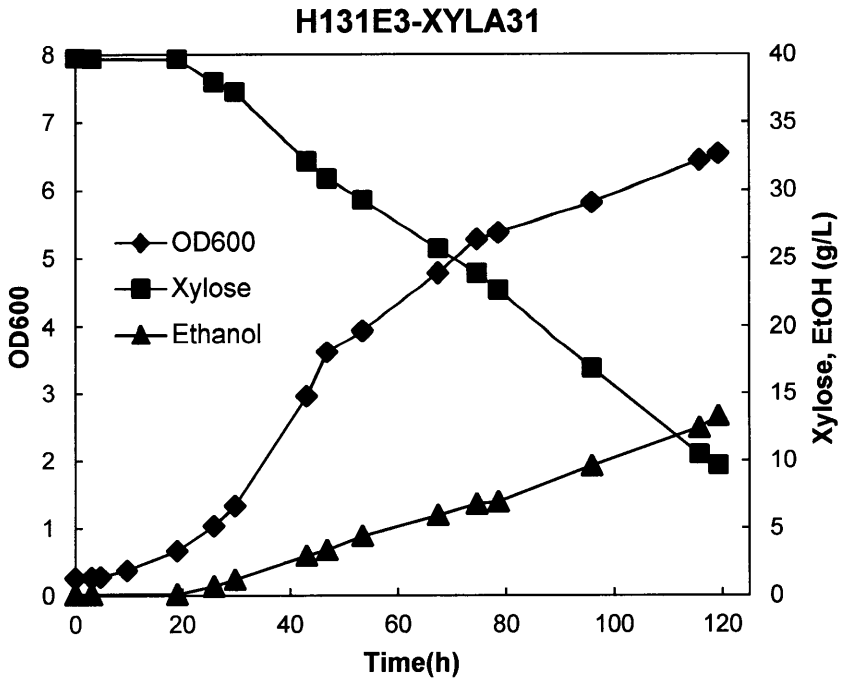


Figure 3-5. Fermentation profiles of H131E3-XYLA31

Table 3-3. Characteristics of H131E3-XYLA31 fermentation

		Growth phase	Overall
μ_{max}	h^{-1}	0.061±0.000	/
Δt	h	76.90	116.25
Δ[Xylose]	g/L	22.79	30.04
Δ[Ethanol]	g/L	9.51	13.20
Δ[DCW]	g/L	1.375	1.673
Δ[Glycerol]	g/L	0.934	1.272
$OD_{average}$		3.793	3.811
$DCW_{average}$	g/L	1.010	1.016
q_{xylose}	$g \cdot L^{-1} h^{-1}$	0.296	0.258
	$g \cdot g^{-1} h^{-1}$	0.293	0.254
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	0.124	0.114
	$g \cdot g^{-1} h^{-1}$	0.122	0.112
EtOH yield	g/g	0.417	0.440
Biomass yield	g/g	0.060	0.056
Glycerol yield	g/g	0.041	0.042

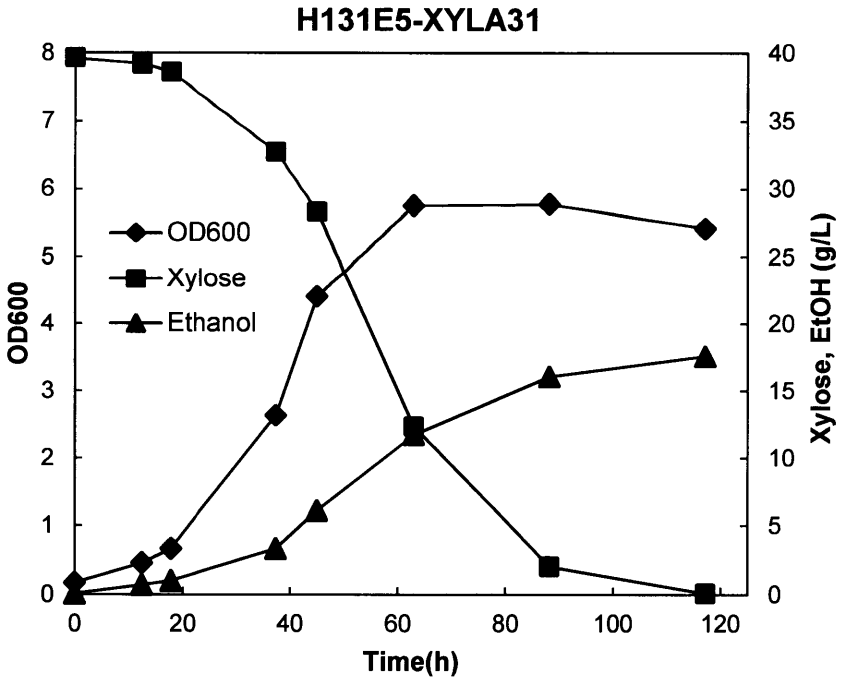


Figure 3-6. Fermentation profiles of H131E5-XYLA31

Table 3-4. Characteristics of H131E5-XYLA31 fermentation

		Growth phase	Overall
μ_{max}	h^{-1}	0.073±0.002	/
Δt	h	50.55	88.15
Δ[Xylose]	g/L	26.85	37.61
Δ[Ethanol]	g/L	11.02	15.97
Δ[DCW]	g/L	1.490	1.495
Δ[Glycerol]	g/L	1.228	1.949
$OD_{average}$		3.043	3.439
$DCW_{average}$	g/L	0.811	0.917
q_{xylose}	$g \cdot L^{-1} h^{-1}$	0.531	0.427
	$g \cdot g^{-1} h^{-1}$	0.655	0.465
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	0.218	0.181
	$g \cdot g^{-1} h^{-1}$	0.269	0.198
EtOH yield	g/g	0.410	0.425
Biomass yield	g/g	0.055	0.040
Glycerol yield	g/g	0.046	0.052

The evolved strain H131E5-XYLA31 displayed an enhanced anaerobic growth rate on xylose 20% higher than H131E3-XYLA31. However, the final biomass concentrations for both strains were very similar. The data from the growth phase and the overall fermentation process were used to calculate xylose consumption, ethanol production and yield. The xylose consumption and ethanol production rates of strain H131E5-XYLA31 were about 2-fold compared with H131E3-XYLA31, while the ethanol yields were similar (0.42-0.44 g/g xylose) for the two strains. Considering the slower xylose consumption and similar biomass accumulation of the micro-aerobically evolved H131E3-XYLA31 to compare with H131E5-XYLA31, H131E3-XYLA31 had about a 60% higher biomass yield during the growth phase, probably due to the emphasis on growth during the evolution with oxygen. It took about 120 hours for H131E5-XYLA31 to consume all the xylose, and for H131E3-XYLA31 the time required was more than 140 hours (data not shown). Both strains had similar glycerol yields (0.04-0.05 g/g xylose) and low xylitol yields (about 0.005 g/g xylose).

3.3.3 Evolution by xylose-limited chemostat cultivation

In the third stage of evolution, the selected strain H131E5-XYLA31 underwent continuous culture under anaerobic conditions using xylose as the limiting carbon source. After the batch reached the stationary phase, chemostat cultivation was started by feeding YNBX to the fermentor at an initial dilution rate of 0.02 h^{-1}. The dilution rate was increased stepwise (0.02 h^{-1}) whenever the culture reached a steady state, which is defined as a stable cell (< 5% change of OD_{600}) concentration for 48 hours or 3× the retention time, whichever came first. After approximately 30 generations, the dilution rate had been increased from 0.02 to 0.08 h^{-1} and the cell density OD_{600} was lowered to 2~3, as compared to initial values of 4~5 in the beginning. The OD_{600} was maintained around 2 while further increasing the dilution rate for another 30 generations, until the dilution rate reached 0.12 h^{-1}.

In order to improve the uptake of xylose at a lower titer, the xylose concentration of the feed was reduced from 20 g/L to 15 g/L for a period equivalent to 20 generations, and then to 10 g/L for another 20 generations. However, under 10 g/L xylose and a dilution rate of 0.12 h^{-1}, the cell density dropped significantly, implying that the low extracellular xylose concentration at this level seriously impeded its transport. Therefore, further engineering on transporters may be necessary to improve transport of xylose at a low extracellular concentration.

In the end, the whole cell population was analyzed at a feed xylose concentration of 15 g/L and dilution rate of 0.148 h^{-1} (the maximum dilution rate that could be used to maintain OD>2). After the continuous culture reached a steady state at OD_{600} about 2.2, the xylose consumption rate was 1.568 $g \cdot g^{-1} h^{-1}$, with an ethanol yield of 0.438 g/g xylose

(Table 3-5). A single strain was then isolated and named H131E8-XYLA31, which was characterized in anaerobic batch cultivation for comparison with its parent strains.

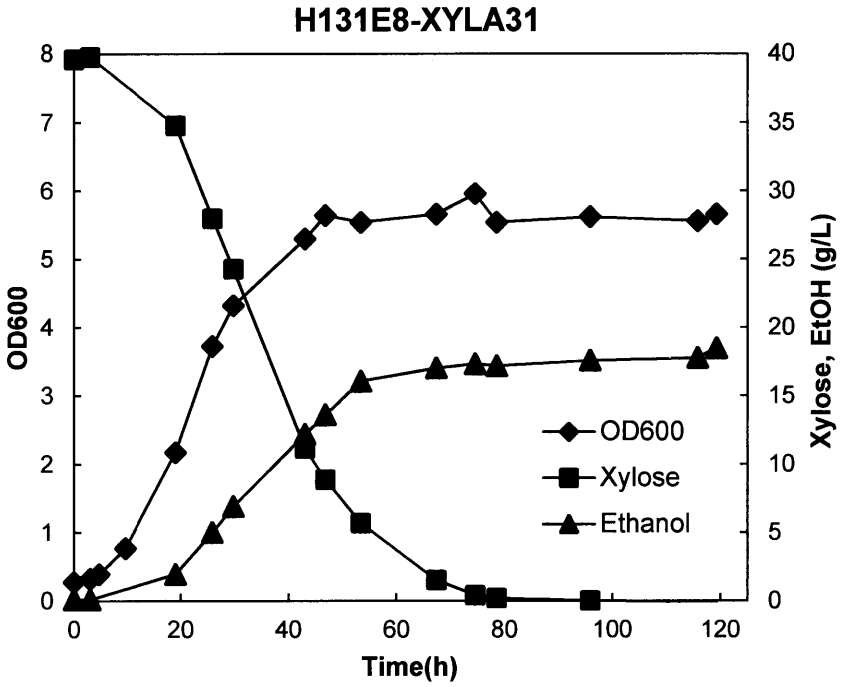


Figure 3-7. Fermentation profiles of H131E8-XYLA31

Table 3-5. Characteristics of H131E8-XYLA31 fermentation

		Chemostat	Growth phase	Overall
μ_{max}	h^{-1}	0.148	0.120±0.004	/
Δt	h	/	40.05	71.60
Δ[Xylose]	g/L	6.206	28.62	39.35
Δ[Ethanol]	g/L	2.719	12.10	17.22
Δ[DCW]	g/L	<0.03	1.327	1.503
Δ[Glycerol]	g/L	N/A	1.057	1.565
$OD_{average}$		2.190	2.926	4.117
$DCW_{average}$	g/L	0.584	0.78	1.098
q_{xylose}	$g \cdot L^{-1} h^{-1}$	0.916	0.715	0.550
	$g \cdot g^{-1} h^{-1}$	1.568	0.916	0.501
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	0.401	0.302	0.240
	$g \cdot g^{-1} h^{-1}$	0.687	0.387	0.219
EtOH yield	g/g	0.438	0.423	0.438
Biomass yield	g/g	0.094	0.046	0.038
Glycerol yield	g/g	N/A	0.037	0.040

In the batch fermentation, the xylose consumption rate of strain H131E8-A31 was about 40% higher than that of H131E5-A31, while the ethanol and glycerol yields remained about the same. It took about 80 hours for H131E8-A31 to consume all the xylose, significantly faster than for H131E5-A31 and H131E3-A31.

3.3.4 Discussion

In the present work, evolutionary engineering was used to improve growth and fermentation of a recombinant *S. cerevisiae* strain on xylose. The strain used to initiate evolution contains the XI-XK xylose-metabolizing pathway, as well as over-expression of the PPP. Strain performance improvements were clearly demonstrated during the three-stage evolution. Both the anaerobic growth rate and the xylose consumption rate significantly improved in the final evolved strain H131E8-XYLA31.

In the three-stage evolution process, different aspects were focused on during each stage. The first aerobic stage emphasized cell growth with a relatively rich medium (SDX). The biomass yield was as high as ~0.5 g DCW/g xylose, while the xylose uptake rate was not high and the ethanol production was rather low. This stage aimed to validate the xylose metabolic pathway, as well as to resolve any possible inhibition of growth caused by the non-preferred carbon source xylose.

The purpose of the second stage of micro-aerobic/anaerobic sequential batch cultivation was to enhance xylose assimilation, since fermentation is the major way that cells utilize xylose under such oxygen-limited conditions, and the cell growth should highly correlate with xylose consumption. To compare with continuous cultivation, each sequential batch underwent various growth phases from low initial OD_{600} to saturation, which potentially allowed selection for a shorter lag phase, better growth in the stationary

phase, or higher final biomass accumulation. Unfortunately, none of the above phenotypes was observed in the strain H131E5-XYLA31 after the evolution, to compare with the strain H131E3-XYLA31.

To further accelerate xylose assimilation, the continuous chemostat evolution was carried out at a situation of relatively low cell density (OD_{600} 2~3) and moderate residue xylose concentration (9~15 g/L), which was equivalent to the fast exponential growth phase in batch culture. Therefore, a remarkable improvement of the growth rate of H131E8-XYLA31 was achieved during the chemostat evolution. The lag-phase of H131E8-XYLA31 on xylose was also shorter than that of H131E3-XYLA31 and H131E5-XYLA31; this relative brevity partially contributes to the earlier finish of fermentation (~80 hr) than the other two strains (~120 hr and more). In the chemostat cultivation, the xylose-limited conditions represent a strong selective pressure for improved uptake kinetics of the growth-limiting substrate xylose. The cell populations in chemostat cultures tend to evolve towards a higher affinity for uptake of xylose; this tendency was reflected by the decrease in the residual xylose concentration even when the dilution rate elevated stepwise.

In the latter phase of the evolution, the feed xylose concentration was reduced from 20 to 15 and eventually 10 g/L in order to improve its uptake rate at a low concentration, which helps to reduce the residual xylose in the end of each batch. As a result, the residual xylose concentration decreased to below 5 g/L when 10 g/L was used in the feed, which greatly slowed down the specific xylose consumption rate and also lowered the cell density. Therefore, H131E8-XYLA31 was isolated at a feed xylose concentration of 15 g/L and a dilution rate of 0.15 h^{-1}. In the batch fermentation, it took

about 35 hours for H131E8-XYLA31 to use up the last 10 g/L xylose, which is better than H131E5-XYLA31 (~50 hours) and H131E3-XYLA31 (not used up till the end of fermentation). This shorter period is likely due to improvement of xylose transport, although this is not confirmed experimentally; it suggests that extensive chemostat evolution under low feed xylose concentration (and subsequently low residual xylose concentration) may be necessary to further improve the transport activity.

As summarized in Table 3-6, the final evolved strain, H131E8-XYLA31, showed a maximum anaerobic growth rate of 0.120 h^{-1} as well as xylose consumption rate of 0.916 $g \cdot g^{-1} h^{-1}$ in a batch cultivation. The data were taken from the growth phase to reflect the maximum capacity of each strain in terms of xylose consumption and ethanol production and also to be comparable with the data in the literature. The performance of H131E8-XYLA31 is similar to that of the strains RWB 217 and RWB 218, which were constructed and evolved using similar strategies. In batch fermentation, H131E8-XYLA31 had the same growth rate but a 25% lower xylose consumption and ethanol productivity, with similar ethanol and xylitol yields. In chemostat cultivation, however, the best instantaneous specific growth rate (0.148 h^{-1}) and xylose consumption rate (1.568 $g \cdot g^{-1} h^{-1}$) were achieved, and both were about 25% higher than the best reported xylose-fermenting strain, RWB218. Moreover, the strain H131E8-XYLA31 had a much lower biomass yield (0.046 g/g) than RWB218 (0.100 $\pm$ 0.003 g/g) [21], suggesting inadequate strain construction or under-optimized fermentation conditions, which will be discussed in Chapter 6.

Table 3-6. Characteristics of xylose-fermenting strains

	Strain description	Fermentation condition	Yields g/g		Ethanol productivity $g \cdot g^{-1} h^{-1}$	Xylose cons. rate $g \cdot g^{-1} h^{-1}$	μ_{max} h^{-1}
			Ethanol	Xylitol			
-XYLA31	*XYLA, PsXYL3, PsTAL1, TKL1, RPE1, RKI1*	Aerobic batch, SDX, 2% xylose	N/A*	N/A	N/A	N/A	0.031±0.02
E1-XYLA31	Selection of H131-A31, aerobic sequential batch	Micro-aerobic batch, SDX, 2% xylose	0.273	<0.01	0.041	0.148	0.197±0.00
E3-XYLA31	Selection of H131-A31, micro-aerobic sequential batch	Anaerobic batch, 2×YNB, 4% xylose	0.417	<0.01	0.122	0.293	0.061±0.00
E5-XYLA31	Selection of H131E3-A31, anaerobic sequential batch	Anaerobic batch, 2×YNB, 4% xylose	0.410	<0.01	0.269	0.655	0.073±0.00
E8-XYLA31	Selection of H131E5-A31, xylose-limited anaerobic chemostat	Anaerobic batch, 2×YNB, 4% xylose	0.423	<0.01	0.387	0.916	0.120±0.00
		Anaerobic chemostat, 1×YNB, 1.5% xylose	0.438	<0.01	0.687	1.568	0.148
202-AFX[19]	*Piromyces XYLA* evolved isolate	Anaerobic batch, synthetic medium	0.42	0.021	0.14	0.34	0.03
217[20]	*Piromyces XYLA, XKS1, TAL1, TKL1, RPE1, RKI1, Δgre3*	Anaerobic batch, synthetic medium	0.43	0.003	0.46	1.06	0.09
218[3, 21]	Selection of RWB 217	Anaerobic batch, synthetic medium	0.41	0.001	0.49	0.9-1.2	0.12

*N/A = not tested

Interestingly, two independent metabolic engineering attempts to achieve xylose-fermenting strains using similar strain construction and evolution strategies resulting in strains with similar performance. This result implies that a limit of the xylose metabolism may have been reached in the current strains, while some major change of a regulatory mechanism or introduction of essential heterologous genes (like a heterologous xylose transporter) may be necessary for further improvement.

3.4 Conclusions

In this chapter the evolutionary engineering approaches were put into practice on the metabolically engineered *S. cerevisiae* strain H131-XYLA31 for improved xylose utilization. The multiple-stage evolution includes aerobic to oxygen-limited sequential batch cultivation and chemostat cultivation with xylose as the limiting carbon source. Both the anaerobic growth rate and the xylose consumption rate were significantly improved after the evolution. A final single-strain isolate H131E8-XYLA31 rapidly grew on and fermented xylose anaerobically in synthetic medium, with a specific growth rate of 0.12 h^{-1} and a specific xylose consumption rate exceeding 0.9 $g \cdot g^{-1} h^{-1}$. The fermentation performances were comparable to those of the best reported strain, RWB 218. However, the biomass yield of H131E8-XYLA31 was significantly lower than the peer strain RWB 218 regardless of the similar initial growth rate, suggesting the potential benefit of further metabolic engineering of the strain and optimization of the fermentation conditions.

CHAPTER 4. Inverse Metabolic Engineering for Identification of Key Genotypes in the Evolutionarily Engineered Strains

4.1 Introduction

Upon successful evolutionary engineering of *S. cerevisiae* towards rapid xylose utilization, the resulting strain, H131E8-XYLA31, could be used as the host for further rational metabolic engineering. However, it is of great interest to identify the genotypes responsible for quick xylose utilization, to further improve the strain or transfer them to a production host with other favorable features, which is essentially inverse metabolic engineering [105, 106]. An inverse metabolic engineering approach based on functional complementation and high-throughput screening has been successfully applied to identify valid gene targets for improved xylose assimilation [18].

In this research, a similar strategy of functional complementation was employed on the evolutionarily engineered xylose-utilizing strains to identify genotypes that promote xylose fermentation. This genotype identification was done by transforming an unevolved *S. cerevisiae* strain, H131-XYLA31, with the genomic library of an evolved strain, H131E5-XYLA31, and then screening the resulting library on a xylose medium. A microfluidic sorting system that sensitively detects quick xylose consumption was used to screen the library. The main target identified was the duplication of the gene *XYLA*, which was confirmed by the structural analysis of the screened plasmid and the qPCR analysis of H131E5-XYLA31.

4.2 Materials and Methods

4.2.1 Construction of genomic library

Genomic DNA was isolated from the target *S. cerevisiae* strain and partially digested with restriction enzyme *Sau3A* I. Then, DNA fragments in the size range from 3 to 8 kbp were isolated from an agarose gel after electrophoresis and purified with ethanol precipitation. Shuttle vector pRS415 [99] was used as the backbone and digested with *Sal* I. Both the genomic DNA fragments and the digested vector were treated with the DNA polymerase I Klenow Fragment (3'→5' exo-) and the appropriate dNTPs to reduce the 5' overhang length from 4 to 2 nucleotides to avoid self-ligation of the vector as well as ligation among genomic DNA fragments, as shown in Table 4-1.

Table 4-1. Genomic library cloning strategy

	Recognition sites	After digestion	After Klenow Fragment treatment
Sal I	5'-GTCGAC 3'-CAGCTG	5'-G 3'-CAGCT	5'-G**TC** 3'-CAGCT
Sau3A I	5'-NGATCN 3'-NCTAGN	5'-N 3'-NCTAG	5'-N**GA** 3'-NCTAG

After ligation of the fragments and the pRS415 backbone by T4 ligase, the resulting plasmids were transformed into ElectroMAX DH5α-E (*Invitrogen*) and plated on agar petri dishes with ampicillin. The colony-forming unit (CFU) was counted, and plasmid was extracted for analysis and transformation of yeast.

4.2.2 Microfluidic high-throughput screening

This high-throughput screening platform developed by Dr. Benjamin Wang [107] uses microfluidic droplet emulsion technology to select for strains from libraries created for inverse metabolic engineering applications. Wang's integrated system has the

capability to encapsulate yeast cells in monodispersed nanoliter-volume aqueous droplets surrounded by an immiscible fluorinated oil phase, culture the cells, mix the contents of the droplets containing cells with a fluorescent enzymatic assay, measure the resulting fluorescence, and sort droplets based on that measurement.

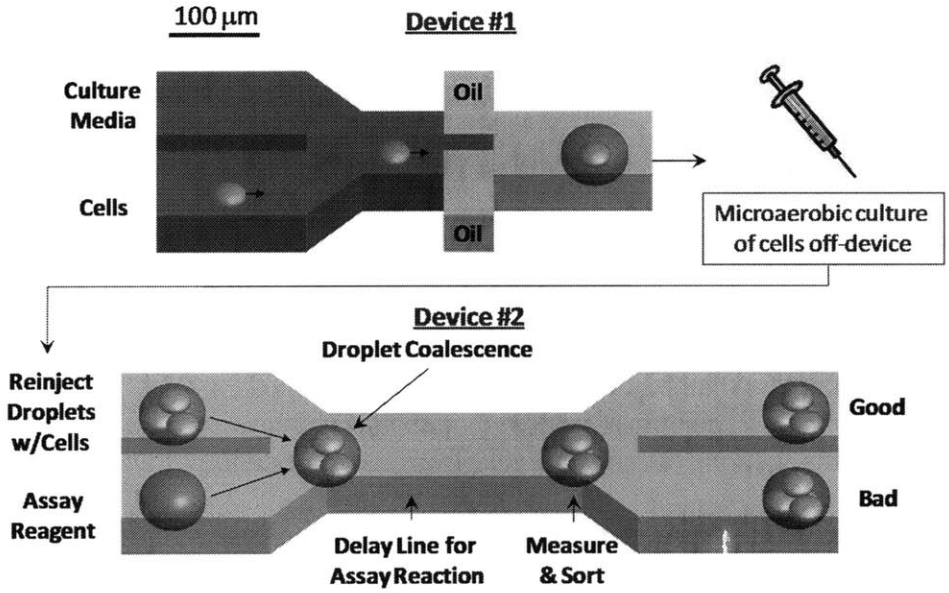


Figure 4-1. Schematic of the microfluidic high-throughput screening platform

In the system shown in Figure 4-1, yeast cells in phosphate buffered saline (PBS) are initially mixed with cell culture medium. Droplets are formed by combining this aqueous stream with two streams containing a fluorinated oil and surfactant mixture. The 0.3nL droplets formed in this device are collected in a syringe, which provides a micro-aerobic environment when capped. The syringe is placed in an incubator for culturing. After being cultured for a predetermined amount of time, droplets from the incubated syringe are re-injected into a second device, where they are combined with another set of droplets containing fluorescent enzymatic assay reagents. After droplet coalescence, the resulting droplets flow through channels for 30 seconds to allow the assay reaction to proceed, as shown in Figure 4-2.

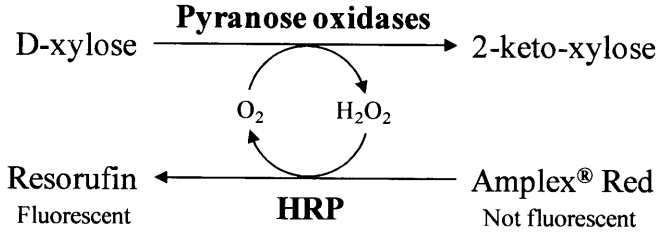


Figure 4-2. Schematic of the xylose assay reactions

The extracellular concentration of the metabolite of interest is determined by measuring the droplet fluorescence with a laser/photomultiplier tube system. Based on this measurement, the droplets are sorted into one of two channels. This system as currently configured can screen approximately 1-2 clones per second so that 10^4 clones can be screened in less than 3 hours.

Note: This section is derived largely from the dissertation of Dr. Benjamin Wang [107]. For more details about the materials and methods please refer to Chapter 2 of that dissertation.

4.2.3 Strains, plasmids, and other materials and methods

The plasmids and strains used in this chapter are summarized in Table 4-2.

Table 4-2. *S. cerevisiae* strains used in this chapter

Plasmids and strains	Characteristics
pRS415	Shuttle vector, CEN, *LEU3*
pRS415-geno	Genomic library of H131E5-A31
pRS415-geno-W2	Isolate from pRS415-geno
E. coli DH5α	Cloning host
H131-XYLA31	H131 (pRS426-*XYLA3*), *Piromyces XYLA, PsXYL3*
H131E5-XYLA31	Evolutionarily engineered strain, isolated from anaerobic sequential batch cultivation
H131-XYLA31-geno	H131-XYLA31 (pRS415-geno)
H131-XYLA31(pRS415)	Control strain of H131-XYLA31-geno

For other materials and methods such as yeast transformation and quantitative PCR, please refer to Chapter 2, Materials and Methods.

4.3 Results

The inverse metabolic engineering approach was attempted by complementing the unevolved *S. cerevisiae* strain of H131-XYLA31 with genomic DNA fragments from the evolved strain of H131E5-XYLA31 to identify factors that promote xylose fermentation, as described in Figure 4-3.

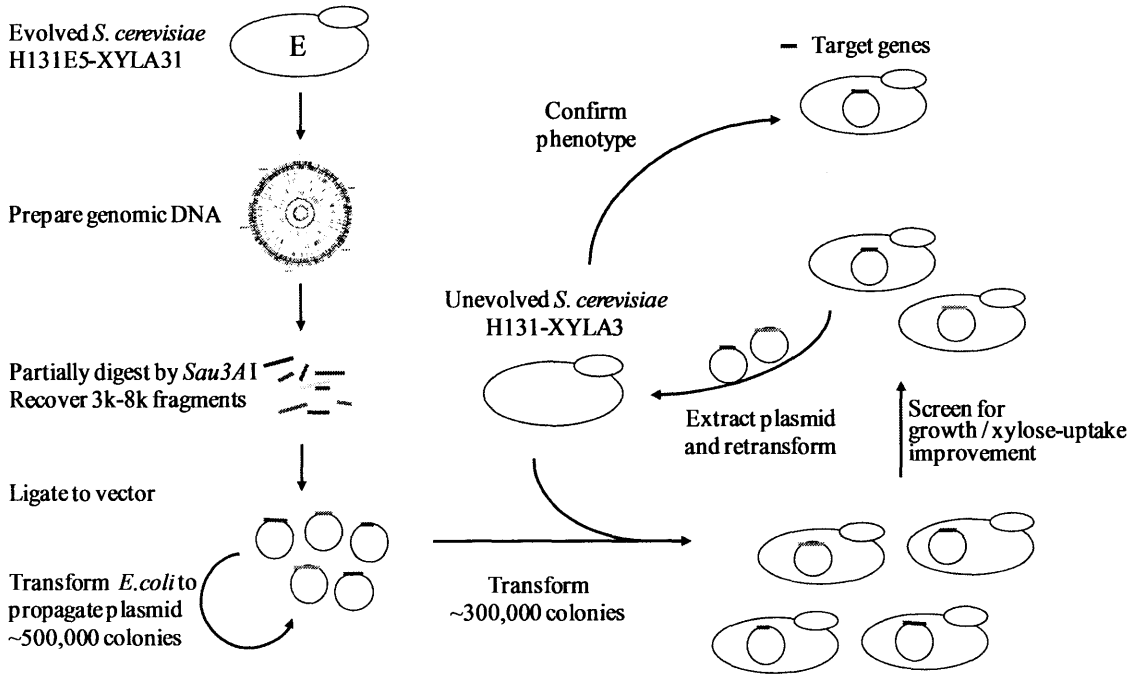


Figure 4-3. Genomic library construction and screening

4.3.1 Genomic library construction

The genomic DNA of *S. cerevisiae* H131E5-A31 was isolated, digested, recovered, modified, and ligated to predigested vector pRS415 to make the genomic

library pRS415-geno. The transformation of pRS415-geno to *E. coli* yielded 5×10^5 CFU on plates.

The quality of the library was assessed by blue-white screening. Based on results of colony-counting after series dilution and replating on an X-Gal plate, about 90% of the colonies contained inserts. The quality of the library was also checked by plasmid digestions, which showed an average insertion size of 3-5k bps (data not shown).

The library was then transformed to H131-A31and plated on an SD medium with 20 g/L glucose and leucine drop-out, to yield about 3×10^5 yeast colonies on plates.

4.3.2 Genomic library screening

Two methods were used for screening of the yeast library H131-XYLA31-geno, including a high-throughput microfluidic screening system, which targets xylose consumption of each single cell, and a screening based on cell growth rate by serial sub-culturing enrichment followed by plating on xylose medium.

The microfluidic screening approach of the genomic library is described in Chapter 3 of Wang's thesis, Screening for High Xylose Consuming Yeast Strains [107]. One round of screening was performed to select only a portion of the droplets, which is about 13% coverage of the library.

For the other approach based on growth rate and colony formation, after 5 rounds of sequential batch cultivation on SDX medium, cells were plated onto SDX agar plates. SDX agar plates tended to dry out at 30°C during incubation over one week. Therefore, for the selection of the library, plates were incubated in a secondary semi-confined container with free water to retain moisture. Only about 50 colonies formed over extended incubation. However, a similar number of colonies grew on the negative control

H131-XYLA31 (pRS415) plate, implying that either the cell grew on another carbon source on the agar plates or other background mutations (evolution) occurred during the extended subculturing. Sixteen larger colonies on the H131-XYLA31-geno plates were picked for further evaluation.

4.3.3 Validation of the selected clones

The 16 selected larger colonies from the growth screening and 8 of the faster xylose-consuming strains isolated from the microfluidic device screening were grown on an SDX medium to verify growth rate, which was followed by the pRS415-geno plasmid isolation and retransformation to H131-XYLA31 to further confirm the improved growth/xylose-assimilating phenotypes.

Five isolates from the microfluidic device screening and one from the growth screening passed the verification. The mutant W2 isolated from the microfluidic screening showed the highest xylose consumption rate.

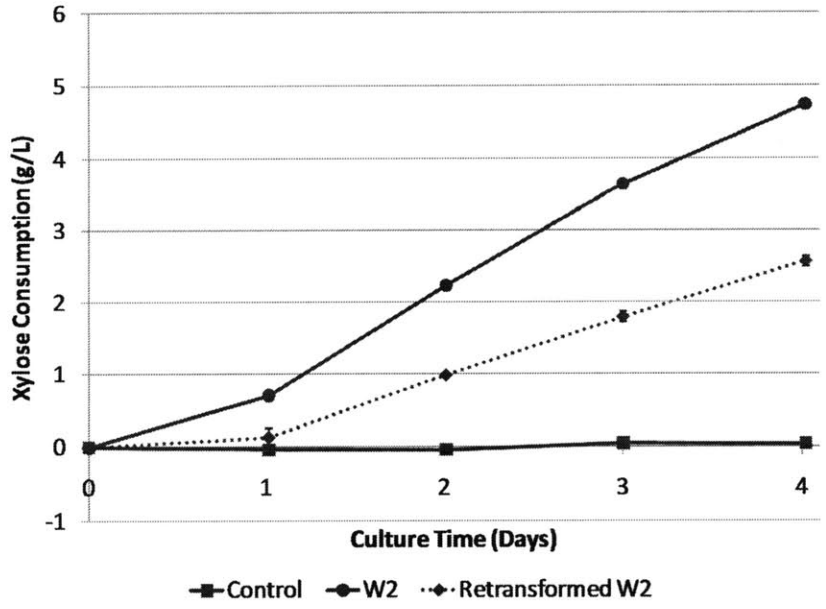


Figure 4-4. Xylose consumption of strain H131-A31 transformed with plasmids [107]

The cumulative xylose consumption in the course of 4 days for the control H131-XYLA31 (pRS415), the mutant H131-XYLA31-geno (W2), and H131-XYLA31 with the retransformed W2 plasmid H131-XYLA31-geno (W2-R) was generated using biological replicates (Figure 4-4 [107]). The retransformed W2-R strain had a xylose consumption of 2.6g/L after 4 days of culturing compared to 4.7g/L for mutant W2, suggesting that a background mutation also occurred in the W2 mutant.

All plasmids were prepared for further analysis of the insertion by DNA sequencing and restriction enzyme digestion.

4.3.4 Analysis of selected mutants

Interestingly, all five plasmids displayed highly similar structures. Here, we show the analysis of one plasmid (pRS415-geno-W2) from the microfluidic screening to demonstrate the structure. DNA sequencing using primers complementary to the plasmid backbone and the *XYLA* gene allowed for the determination that the plasmid pRS415-geno-W2 contained at least one full ORF of the *XYLA* gene plus two truncated fragments of the ORF (Figure 4-5).

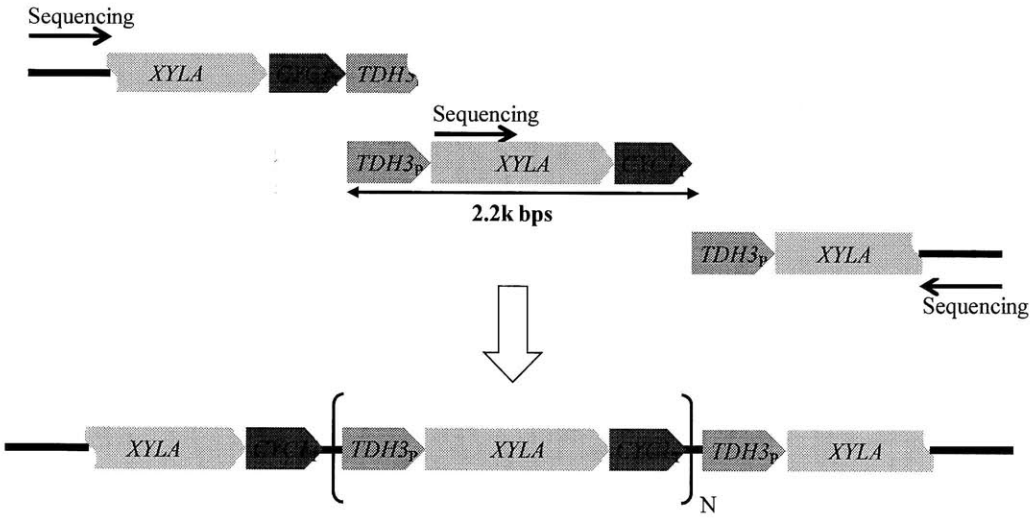


Figure 4-5. Sequencing results of plasmid pRS415-geno-W2

For more details about the plasmid pRS415-geno-W2, the restriction enzyme *Kpn* I was used to digest the plasmid and confirm that the *XYLA* gene construct (*TDH3p-XYLA-CYC1t*) with a size of 2.2 kbps did exist on the plasmid (Figure 4-6 and Table 4-3, Lane 1), and likely more than one copy duplicated since the intensity of the 2.2 kbps band was even higher than that of the 5319 bps band. As a control, digestion analysis for the plasmid pRS415 was also performed using the same restriction enzymes (Figure 4-6, Lanes 2, 4).

To determine the *XYLA* gene copy number on pRS415-geno-W2, an *EcoR* I digest was performed to confirm that the total size of the plasmid was about 14 kbps (Figure 4-6, Lane 3); therefore, three full copies of the *TDH3p-XYLA-CYC1t* construct were calculated by comparing Lane 1 and Lane 3 of Figure 4-6 [(~10000+3962) – (5319+1621+688) bps ≈ 3×2.2 kbps]. The copy number was also verified by analyzing the band intensity of Lane 1 of Figure 4-6 with software ImageJ (NIH, data not shown). Through these analyses, the sequence of pRS415-geno-W2 was proved to have three full *XYLA* gene constructs with truncated constructs flanking on each side (Figure 4-7).

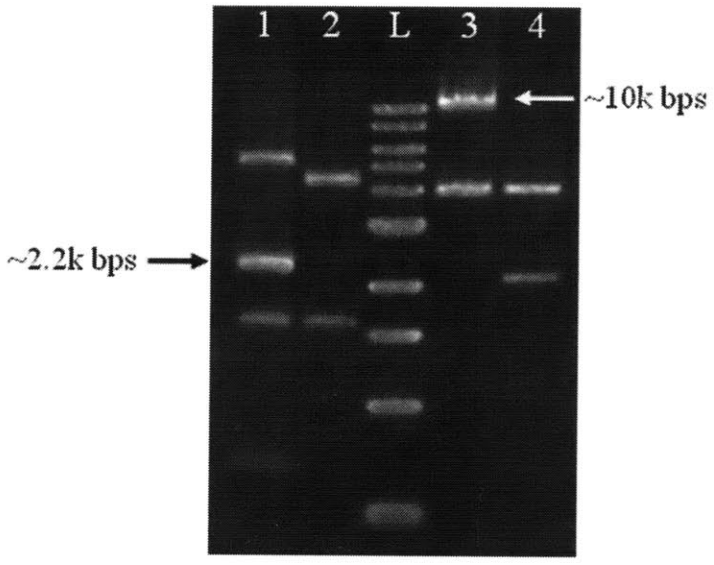


Figure 4-6. Restriction enzymes digestion of plasmid pRS415-geno-W2 and control pRS415

Table 4-3. Restriction enzyme digestion of plasmid pRS415-geno-W2 and control pRS415

Lane	1	2	3	4	5
Restriction Enzyme	KpnI		\	EcoRI	
Plasmid	pRS415-geno-W2	pRS415	DNA ladder	pRS415-geno-W2	pRS415
Fragments (bps)	5319 2264 1621 688	4400 1621	10000 8000 6000 5000 4000 3000 2000 1500 1000 500	10,458 3962	3962 2059

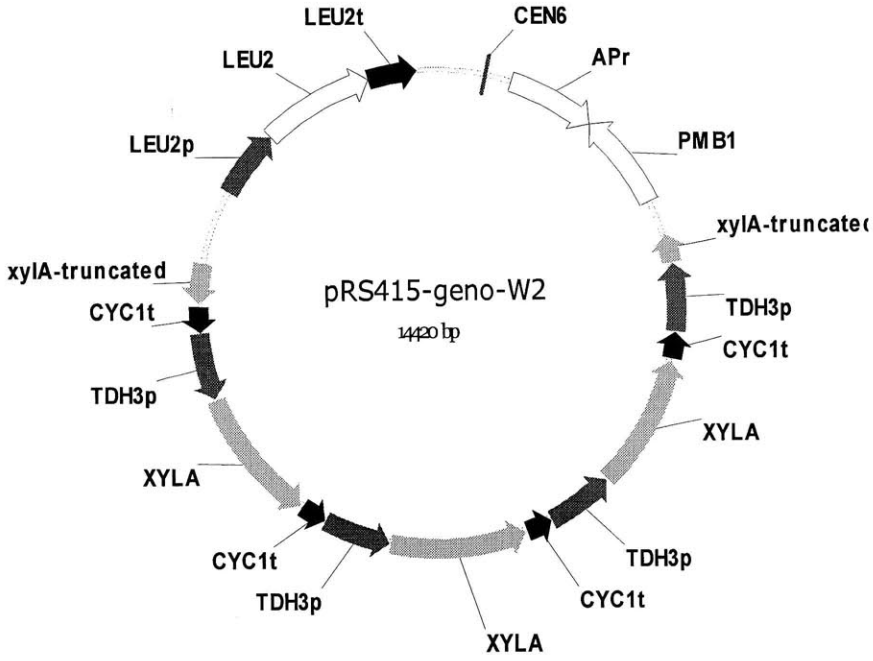


Figure 4-7. Map of plasmid pRS415-geno-W2

DNA sequencing also identified a S19Y (serine to tyrosine) point mutation in the xylose isomerase enzyme; however, the effect of the point mutation was not confirmed (data not shown).

4.3.5 Copy number of genes from the xylose metabolism pathway

Since pRS415-geno-W2 carries only a fragment of the H131E5-XYLA31 genome, it can be expected that more tandem copies of *XYLA* exist in the H131E5-XYLA31 genome. Also, it is interesting to find the duplication of another key xylose metabolic gene, *XYL3,* which was introduced together with *XYLA*.

The copy numbers of genes *XYLA* and *XYL3* in the unevolved strain H131-XYLA31 as well as the evolved strains H131E5-XYLA31 and H131E8-XYLA31 were evaluated through quantitative PCR. Endogenous *PGK1* was used as a reference during the quantification, since a single copy of *PGK1* per haploid genome can be assumed. The copy number of *XYLA* and *XYL3* per genome was calculated by comparing their relative amounts to the *PGK1* present in the same sample, as shown in Figure 4-8.

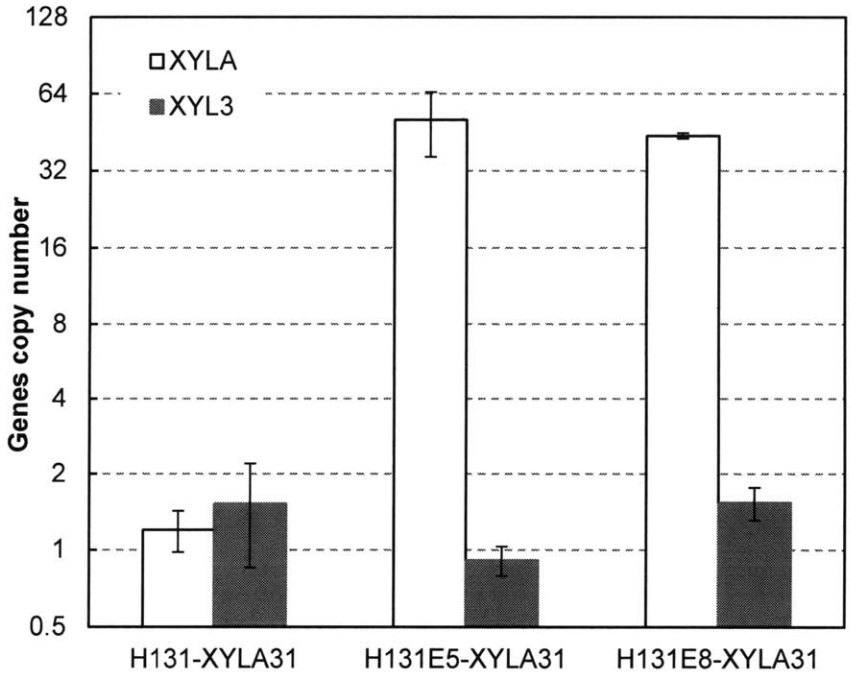


Figure 4-8. Copy number of genes *XYLA* and *XYL3*

For the unevolved strain of H131-XYLA31, 1.20±0.22 and 1.53±0.68 copies of *XYLA* and *XYL3* per genome were found, respectively (Figure 4-8). This result is consistent with the fact that *XYLA* and *XYL3* were carried on the same vector at a 1:1 copy ratio. However, it is interesting to notice that the vector copy number was rather low given that the 2μ plasmid usually keeps 10-20 copies without selective pressure.

For the evolutionarily engineered strains H131E5-XYLA31 and H131E8-XYLA31, a similar high copy number of *XYLA* (50.6±14.4 and 43.7±1.1) was confirmed, while *XYL3* gene's copy remained low for both strains (0.91±0.12 and 1.54±0.23 copies). A higher expression level of *XYLA* by multi-copy is required [20], while moderate expression of *PsXYL3* in a low copy number in addition to endogenous *XKS1* provides enough xylulokinase activity [17]. However, it is interesting to observe the enormous difference between the *XYLA* and *XYL3* copy numbers, given that both genes were cloned on the same vector at a 1:1 copy ratio. The duplication of *XYLA* is further discussed in Chapter 5.

4.4 Discussion

The inverse metabolic engineering approach based on functional complementation enables discovery of dominant genotypes in the evolutionarily engineered strains, providing a good coverage of the genomic library and a high-efficiency screening method. The present study obtained the proper library sizes in both *E. coli* and yeast (10^5-10^6), thanks to the efficient transformation protocols. The genomic library could possibly reveal the dominant genotypes of the target strain of interest. Other recessive genotypes in the strain may require construction of a random knock-out library. Such a library can be easily created for *E. coli* by using transposons; nevertheless,

construction of a yeast knock-out library remains quite tedious, particularly for an evolutionarily engineered strain with a poorly characterized background.

Two approaches were applied for the screening of the library. The microfluidic-based screening system sorts for cells encapsuled in droplets according to xylose consumption of each single cell (or population derived from a single cell in a droplet). The other approach, selection for faster cell growth, is based on the hypothesis that genes that increase xylose assimilation will also promote the growth rate of transformants on a xylose medium. In many other cases the traditional growth-rate screening would work effectively [18, 108]; however, the "good" transformants in the genomic library constructed in the present work displayed only a minor growth advantage, and the screening result is barely satisfactory (1 out of 16 isolates after screening were verified to show fast growth phenotype). This small advantage is largely because the xylose isomerase requires the presence of multiple copies to be potent. The microfluidic-based screening, on the other hand, was proved to be much more sensitive (5 out of 8 isolates after screening were verified to show a fast xylose consumption phenotype and 4 contain meaningful inserts). Nonetheless, only part of the library was sorted due to the limit of the current screening setup. The throughput of the screening needs to be further improved to cover the whole library in a reasonable timeframe.

The plasmid pRS415-geno-W2 isolated from the genomic library contains three full copies of the *XYLA* expression construct plus two truncated copies flanking each side. The insert was confirmed to be digested by *Sau3A* I and ligated in the correct *Sal* I site. Other isolated plasmids contain a similar tandem duplication of *XYLA*, but they vary in copy number (3~4), digestion site (2 in *XYLA* ORF), and orientation of insert. The size of

the insert ranges from 8.4~10.6k bps, and it exceeds the size of DNA collected from the gel (up to 8k bps) for the ligation. Therefore, it is possible that recombination of *XYLA* constructs occurred, resulting in increased copy number during the preculture before the microfluidic screening or sequential culture in the selection process.

From the above observation, it seems possible to acquire higher copies of tandem gene duplication from lower copies. The hypothesis is that the duplication initiated from the unequal crossover of the *TDH3p-XYLA-CYC1t* - *TDH3p-PsXYL3-CYC1t* construct originally on the plasmid pRS426-*XYLA3* and elongated thereafter (Figure 4-9).

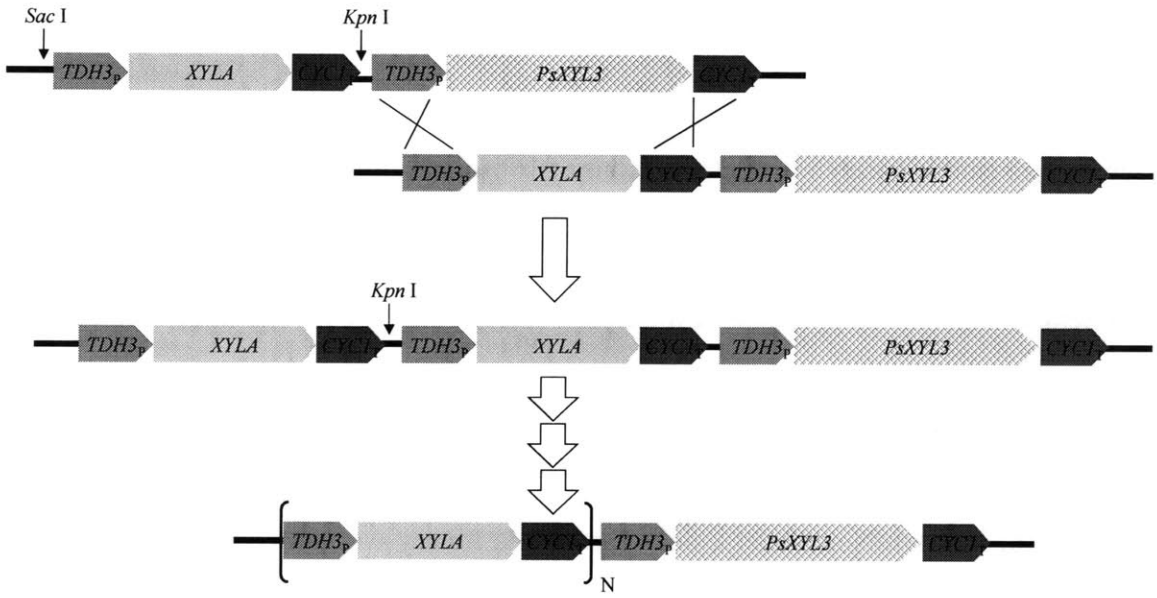


Figure 4-9. Scheme of unequal crossover of pRS426-*XYLA-XYL3*

One piece of evidence supporting this hypothesis was that the *Kpn* I site between the adjacent *XYLA* duplication after evolution was from the original pRS426-*XYLA3*, between the *XYLA* and *XYL3*. Moreover, the recombination could occur more frequently than cell mitosis, since the pRS426-*XYLA3* is a multi-copy plasmid. Nevertheless, repeated attempts to isolate the recombined plasmid pRS426-$(XYLA)_N$-*XYL3* from evolved strains of H131E-XYLA31 were unsuccessful. Further effort to cure the pRS426

plasmid from the H131E-XYLA31 based on the *URA3*/5-FOA counter-selection resulted in a strain without the *URA3* genotype, but retaining xylose-utilizing ability. All the facts implied that the *XYLA* duplication had recombined onto the genome during the evolutionary engineering.

The locations of *XYLA* in the evolved strains were confirmed through PFGE followed by chromosome blotting, as discussed in Chapter 5. The evidence of qPCR results and the tandem gene duplication structure on pRS415-geno-W2 strongly suggested that the duplication and integration of gene *XYLA* codes for XI was a major recombination event during the evolutionary engineering, and a high expression level of heterologous XI is necessary for rapid xylose utilization.

4.5 Conclusions

In this chapter, we demonstrated a successful inverse metabolic engineering approach to identify gene targets responsible for a rapid xylose metabolic phenotype. A genomic library of evolutionary engineered H131E5-XYLA31 was constructed in *E. coli*, then amplified, and finally retransformed to *S. cerevisiae* host for screening. The microfluidic sorting system aiming for quick xylose consumption is much more sensitive than the screening based on growth rate, despite lower coverage of the library. The main target identified by both types of screening is the duplication of *XYLA*, which was confirmed by the structural analysis of the screened mutant and the qPCR of the strain H131E5-XYLA31.

Two obstacles limited our search: first, the difficulty of constructing the yeast knock-out library for identifying recessive phenotypes, and second, the throughput of the microfluidic screening method for a larger library.

CHAPTER 5. Biochemical and Genetic Characterization of the Evolutionarily Engineered Strains

5.1 Introduction

Evolutionary engineering of *S. cerevisiae* containing the XI-XK xylose metabolic pathway yielded strain H131E8-XYLA31, which quickly ferments xylose. In addition, the inverse metabolic engineering approach was successfully employed to identify the tandem *XYLA* duplication as a key genotype in the evolved strains. However, the genotype of the evolved strains still remains unclear and more in-depth knowledge is needed for interpreting and further improving the strain. This chapter describes the characterizations of the engineered xylose-utilizing strains on the genomic, transcriptional, and biochemical levels. In particular, the duplication process of *XYLA* was verified since the method can be possibly used for over-expression of other genes of interest.

The expansion of *XYLA* and consequently high expression level of XI were confirmed by qPCR, blotting techniques, and enzyme activity assay. In the series of evolved H131E-XYLA31 strains, *XYLA* was proved to integrate into the chromosomes and exist in tandem gene duplication (TGD) structure. The evolution process was successfully reproduced, and *XYLA* expansion could originate from different initial *XYLA* constructions. The location-specific *XYLA* integration and duplication could also be achieved. The XK (*XYL3*) form *P. stipitis* could be efficiently expressed in *S. cerevisiae* at low copy number. Mutations on *GRE3* that potentially abolish XR activity were discovered in the evolutionarily engineered strains.

5.2 Materials and Methods

5.2.1 Southern blot

7) Restriction digestion of genomic DNA.

The gDNA was treated as follows. The 10-μg genomic DNA was digested in 400 μl total volume with 3 units of restriction enzyme (*EcoR* I or *Xba* I) per μg DNA at 37°C overnight. Then 0.1 volume 3 M sodium acetate (pH 5.2) was added and mixed well. Next 1 volume ice-cold 100% isoproponal was added. The solution was precipitated at -80°C for 30 minutes and spun in a microfuge at high speed for 20 minutes. After being aspirated, the DNA pellet was washed with 70% ethanol and spun in a microfuge at high speed for 10 minutes. The pellet was aspirated and air-dried (or Vacufuged) for 5 minutes. Finally, it was resuspended in 15 μl 10mM Tris pH 7.5 and 3 μl 6 × DNA-loading buffer was added.

8) Agarose gel electrophoresis

Samples and marker were loaded into 0.8% agarose gel in TAE. They were electrophoresed at 4 Volts/cm for 2-4 hours and then stained by SYBR gold and photographed with a fluorescent ruler.

9) Southern transfer

Denaturation: gel was soaked in two gel volumes of 1.5 M NaCl / 0.5N NaOH for 2 ×15 minutes. Neutralization: gel was soaked in two gel volumes of 1.5 M NaCl / 1M Tris-HCl (pH7.5) for 2 ×15 minutes, and then in two gel volumes of transfer buffer (10 × SSC, Saline-Sodium Citrate, 1.5 M sodium chloride and 150 mM trisodium citrate, pH 7.0) for 15 minutes.

Transfer was set up in a large electrophoresis tray with transfer buffer (10 × SSC). The following items (gel-sized, saturated with transfer buffer) were placed on the middle support: Wick (3mm paper); 3 pieces of Whatman 3MM paper; Gel (upside down); Hybond-N+ membrane (GE); 3 pieces of Whatman 3MM paper; dry 5-cm stack of paper towels; glass plate with weight (~500 g). DNA transfer was carried out overnight.

The next day, the blotting material was taken off and the position of the wells was marked with a very soft lead pencil. The membrane was placed on a wet paper towel (with 2 × SSC) of equal size. The DNA was immobilized by using a Stratalinker® UV Crosslinker (auto crosslink mode).

10) Hybridization

The membrane was placed in a hybridization bottle with 25-ml Church buffer (1% BSA, 1 mM EDTA, 0.5 M Na-PO_4 pH 7.2, 7% SDS), and then incubated in a hybridization oven at 68°C for 1-2 hours. DNA probes labeled by [32P]dCTP were prepared using the Prime-It II Random Primer Labeling Kit (Stratagene, La Jolla, CA). The denatured probe was added to the hybridization solution. The solution was incubated in a hybridization oven at 68°C overnight.

11) Wash

The membrane was soaked 2-3 times for 30 minutes each with 100 ml of wash buffer (0.2 × Church buffer, 0.1% SDS) in a hybridization oven at 68°C, then for 15 minutes in 100 ml of wash buffer at room temperature.

12) Detection

After washing, the membrane was blotted with filter paper (Whatman 3MM) to remove most of the excess moisture. Moist blots were wrapped in plastic wrap prior to

autoradiography. The blots were exposed to a storage phosphor screen (GE) for 1 to 3 days and scan with a Typhoon™ 9400 scanner (GE).

5.2.2 Northern blot

RNA gel electrophoresis was modified from Molecular Cloning [100]. The 4uL RNA, 1uL 0.2M Na-PO_4 (pH 7.0), 10uL DMSO, and 3.75uL 6M (40%) Glyoxal were mixed and incubated at 50 degrees for one hour to denature the RNA. The gel box and comb were treated with 3% H_2O_2 and 0.8% agarose gel in 10mM Na-PO_4 (pH 7.0 in DEPC H_2O). The loading buffer was added to the samples and mixed well. Samples were loaded and run at 4V/cm until bromphenol blue had migrated approximately 8 cms (approximately 3 hours). The neutral transfer and subsequent steps were the same as with the Southern blot.

5.2.3 Pulsed field gel electrophoresis (PFGE) and chromosome blot

Chromosomal DNA was prepared as described by Carle and Olson [109]. Chromosomes were separated on a 1% agarose gel (Bio-Rad) in a 0.5× TBE buffer at 6V/cm at 14 °C, using a Bio-Rad CHEF-DIII mapper apparatus. The following migration conditions were used: a pulse time of 60 s for 15 hours and a pulse time of 90 s for 9 hours with an angle of 120°. The gel was stained with ethidium bromide to identify the chromosomal pattern specific to each strain.

The PFGE gel was treated with UV irradiation using Stratalinker® UV Crosslinker (auto crosslink mode). The subsequent steps for the chromosome blot were the same as for the Southern blot.

5.2.4 XI activity assay

S. cerevisiae was grown to an exponential phase in YNBG or YNBX medium. Cells were harvested by centrifugation for 10 min at 8,000×g and 4°C. The pellet was washed twice with chilled washing buffer (10 mM phosphate buffer, 2 mM EDTA, pH 7.5) and suspended in chilled extraction buffer (100 mM phosphate buffer, 2mM $MgCl_2$, 1mM dithiothreitol, pH 7.5). The suspended cells were mixed with 0.5-mm glass beads (Sigma), vortexed at maximum rate for 1 minute, and then cooled on ice for a similar period. This procedure was repeated for up to 10 min. The glass beads and cellular debris were removed by centrifugation for 15 min at 15,000×g and 4°C, and the crude extract collected was used for the enzyme assay. The total protein concentration in cell extracts was determined using the Pierce Coomassie Protein Assay Reagent Kit (Pierce Biotechnology, Rockford, IL) with bovine serum albumin as the standard.

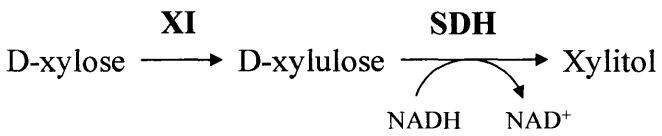


Figure 5-1. Xylose isomerase assay

Xylose isomerase activity in cell extracts was determined spectrophotometrically using a method modified from Kersters-Hiderson et al. [110]: the assay mixture (1 ml) containing 100 mM Tris-HCl buffer (pH 7.5), 10 mM $MgCl_2$, 0.2 mM NADH, 2 U sorbitol dehydrogenase (Roche), and 20 µl of cell extract was equilibrated at 30°C for 5 min. The reaction was started by the addition of D-xylose to a final concentration of 200 mM. The assay was performed at 30°C for 10 min using an Ultrospec 2100 pro UV/Visible Spectrophotometer (GE Healthcare Biosciences). An extinction coefficient at 340 nm of NADH of 6.3 $mM^{-1}cm$- was used to calculate specific activity. One unit of

xylose isomerase activity was defined as the amount of enzyme required to produce 1 μM of xylulose per minute under the assay conditions.

5.2.5 XK activity assay

Xylulokinase activity was measured according to the method of Shamanna and Sanderson [111].

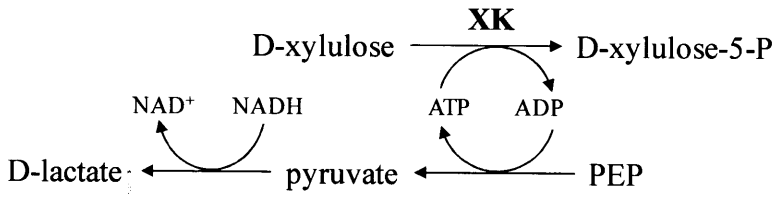


Figure 5-2. Xylulokinase assay

Cell extraction and protein concentration determination were done as described above in the XI assay. The assay mixture (1 ml) containing 20mM tris-HCl (pH7.5), 2mM ATP, 0.2mM PEP, 3mM Glutathione, 0.1mM NADH, 2mM $MgCl_2$, 5mM KCN (NADH oxidase inhibitor), 5mM NaF (ATPase inhibitor), 150mg/L LDH-PK (~30U for each enzyme), and 20uL cell extract was equilibrated at 30°C for 5 min. The reaction was started by the addition of D-xylulose to a final concentration of 2 mM. The assay was performed at 30°C for 10 min using an Ultrospec 2100 pro UV/Visible Spectrophotometer (GE Healthcare Biosciences). An extinction coefficient at 340 nm of NADH of 6.3 $mM^{-1}cm$- was used to calculate specific activity. One unit of xylulokinase activity is defined as the amount of enzyme that phosphorylates 1 μmol of xylulose per minute under the assay conditions.

5.2.6 Plasmids, strains and other materials and methods

The strains and plasmids used in this chapter are summarized in Table 5-1.

Table 5-1. Strains and plasmids used in this chapter

Strains and plasmids	Characteristics
***S. cerevisiae* strains**	
H131	BF264-15Dau , *MATa, leu2-3,112, ura3, arg4, TRP1::RKI1- RPE1, HIS2::TKL1, ADE1::PsTAL1*
H131E1-XYLA31	H131-XYLA31, isolated from aerobic sequential batch cultivation
H131E3-XYLA31	H131-XYLA31, isolated from micro-aerobic sequential batch cultivation
H131E5-XYLA31	H131-XYLA31, isolated from anaerobic sequential batch cultivation
H131E8-XYLA31	H131-XYLA31, isolated from anaerobic chemostat
H131-LEU2	H131, *leu2-3*complemented by pRS405 containing *LEU2*
H131-gre3Δ	H131, *gre3::LEU2*
H142	H131, *ARG4::HXT7p-PsXYL3-CYC1t*
H142-p416A	H142 (pRS416 $HXT7_{Tr}$-XYLA)
H142-p416A2	H142 [pRS416 $HXT7_{Tr}$-$(XYLA)_2$]
H142-p426A	H142 (pRS426GPD-XYLA)
H142-p405A2	H142 [pRS405-$(XYLA)_2$]
H142-p406A2	H142 [pRS406-$(XYLA)_2$]
H142E-p416A	H142-p416A, isolated from aerobic sequential batch cultivation
H142E-p416A2	H142-p416A2, isolated from aerobic sequential batch cultivation
H142E-p426A	H142-p426A, isolated from aerobic sequential batch cultivation
H142E-p405A2	H142-p405A2, isolated from aerobic sequential batch cultivation
H142E-p406A2	H142-p406A2, isolated from aerobic sequential batch cultivation
Plasmids	
pRS415GPD-yECitrine	pRS415, *TDH3p- yECitrine-CYC1t*
pRS415TEF-yECitrine	pRS415, *TEF1p- yECitrine-CYC1t*
pRS415PGK1-yECitrine	pRS415, *PGK1p- yECitrine-CYC1t*
pRS415$HXT7_{Tr}$-yECitrine	pRS415, *$HXT7_{Tr}$p - yECitrine-CYC1t*
pUCAR1	pUC19, *Hind*III-*ARG4-Hind*III
pUCAR1-PsXYL3	pUCAR1, *$HXT7_{Tr}$p-PsXYL3-CYC1t*
pRS416 $HXT7_{Tr}$-XYLA	CEN/ARS ori, *URA3*, *TDH3p-XYLA-CYC1t*
pRS416 $HXT7_{Tr}$-$(XYLA)_2$	CEN/ARS ori, *URA3,* 2 copies of *$HXT7_{Tr}$p-XYLA-CYC1t* in tandem
pRS426GPD-XYLA	2μ ori, *URA3*, *TDH3p-XYLA-CYC1t*
pRS405	YIp, *LEU2*
pRS406	YIp, *URA3*
pRS405-$(XYLA)_2$	*LEU2,* 2 copies of *$HXT7_{Tr}$p-XYLA-CYC1t* in tandem
pRS406-$(XYLA)_2$	*URA3,* 2 copies of *$HXT7_{Tr}$p-XYLA-CYC1t* in tandem

Fluorescence of yECitrine was monitored in a Hitachi F2500 Fluorescence Spectrophotometer at excitation/emission wavelengths of 402nm/432nm, using 1-cm cuvettes. For other materials and methods such as cell cultivation, yeast transformation, and quantitative PCR, please refer to previous chapters.

5.3 Results and Discussion

5.3.1 Verification of tandem gene duplication of *XYLA*

In order to test the hypothesis of unequal crossover, which possibly led to the tandem *XYLA* duplication, the XI-XK pathway was reconstructed in the PPP over-expressing host strain, H131. To achieve a moderate activity of xylulokinase, one copy of *PsXYL3* was integrated into the genome. The *TDH3* promoter for the *XYL3* gene was replaced with a constitutive truncated *HXT7* promoter [112] to avoid further recombination with the *TDH3p-XYLA-CYC1t*. Carried on a low copy CEN/ARS plasmid, the $HXT7_{Tr}$ promoter was tested using codon-optimized GFP yECitrine [113] as a reporter in the host H131E3-XYLA31.

As shown in Figure 5-3, the $HXT7_{Tr}$ promoter showed an expression level similar to that of the *TDH3* promoter and stronger than the other two constitutive promoters of *TEF1* and *PGK1*. The $HXT7_{Tr}$-*p-PsXYL3-CYC1t* construct was then cloned into pUCAR1 at the *EcoR* I- *BamH* I site, and the resulting pUCAR1-*PsXYL3* was linearized with *Nhe* I for transformation of H131, resulting in strain H142 (Table 4-2).

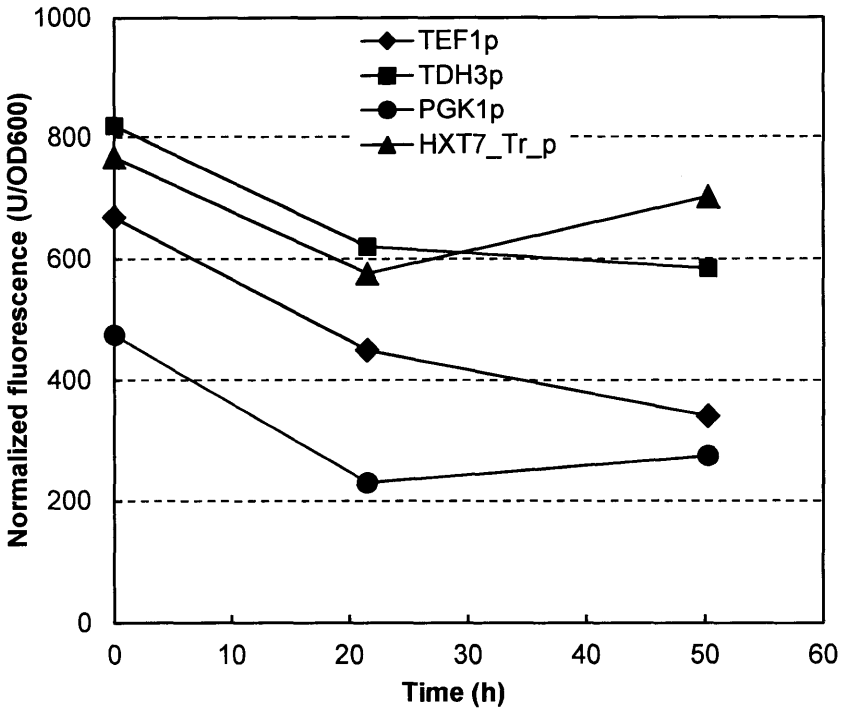


Figure 5-3. Expression of yECitrine under various promoters

For the reconstruction of a xylose-metabolizing strain, H142 containing *XYL3* and PPP genes was transformed with plasmids with *XYLA* (Table 4-2), resulting in the following strains (Table 5-1):

1) H142-p416A, one copy of *XYLA* on a low-copy-number CEN/ARS plasmid;
2) H142-p416A2, two tandem copies of *XYLA* on a low-copy-number CEN/ARS plasmid for verifying unequal crossover on the plasmid;
3) H142-p426A, one copy of *XYLA* on a 2-μ multi-copy plasmid;
4) H142-p405A2, two tandem copies of *XYLA* integrated into the *leu2* locus on chromosome III of H142; and
5) H142-pRS406A2, two tandem copies of *XYLA* integrated into the *ura3* locus on chromosome V of H142.

The resulting strains were evolutionarily engineered using an approach of aerobic sequential batch cultivation similar to that described in Chapter 3, but in a rich medium

(YPX) to accelerate the process. After about 2 weeks, substantial growth on xylose was observed for all tested strains, and the cultivation was switched to the synthetic defined medium SDX for another two weeks, which is equivalent to about 30 generations. The micro-aerobic growth of the resulting strain (named H142E1-vector-XYLA) was compared with that of the previous evolutionarily engineered H131E1-XYLA31 and H131E3-XYLA31, as shown in Figure 5-4.

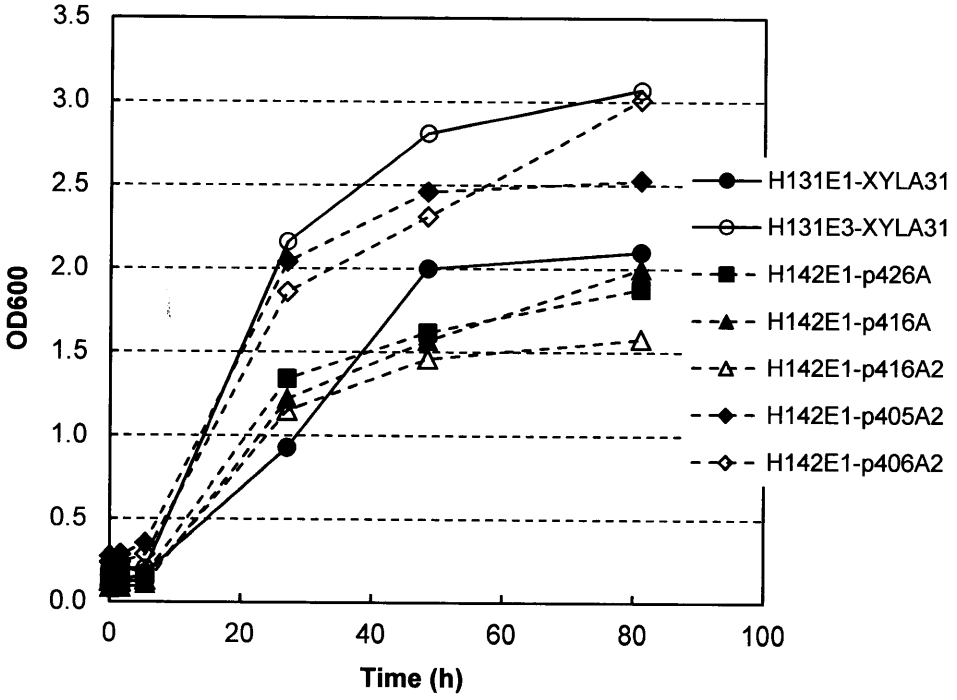


Figure 5-4. Micro-aerobic growth of H131E/H142E strains

All of the newly evolved H142E strains showed a growth rate between that of the aerobically evolved H131E1-XYLA31 and that of the micro-aerobically evolved H131E3-XYLA31, indicating that the new strategy of construction and evolution worked efficiently. However, none of the newly evolved strains out-performed H131E5-XYLA31 (data not shown), implying that extended oxygen-limited evolution is necessary for further improvement of the new strains.

From none of the 3 plasmid-based strains H142E-p416A/p416A2/p426A could the plasmid be segregated after evolution. The observation was the same as for the H131E-XYLA31 strains. Interestingly, the H142E1-p416A containing only one initial copy also got evolved after adaption, suggesting that recombination mechanisms other than the hypothesized unequal crossover may occur.

For the 2 *XYLA* integration strains of H142E-p405A2/p406A2, the locations of the *XYLA* integration were verified by PCR.

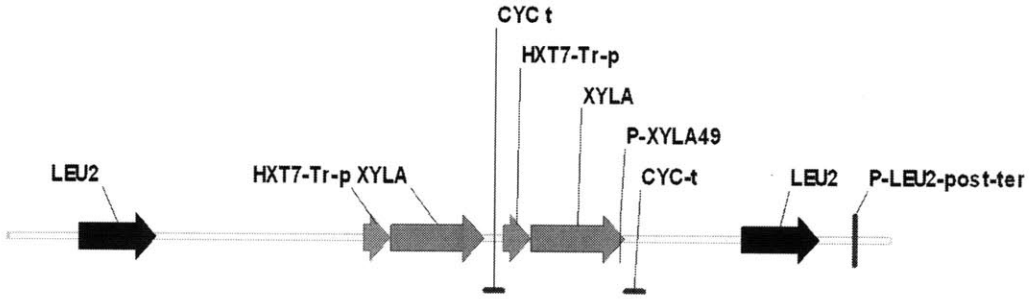


Figure 5-5. $(XYLA)_2$ integration in *LEU2* of H142-p405A2

In H142E-p405A2, primers P-XYL49 (5'-ACAAACCTCTGGTAAGCAGGA ATTGTACGAAGCTATTGTCGCAATGTATCA) in the *XYLA* open reading frame and P-LEU2-post-ter (5'-GCCCATTCTTCCATCAGATTTGGTATTGG) were used to successfully amplify a 3.3-kbps fragment. The size of the PCR product and its sequencing result confirmed the expected location of the *XYLA* integration in *LEU2*.

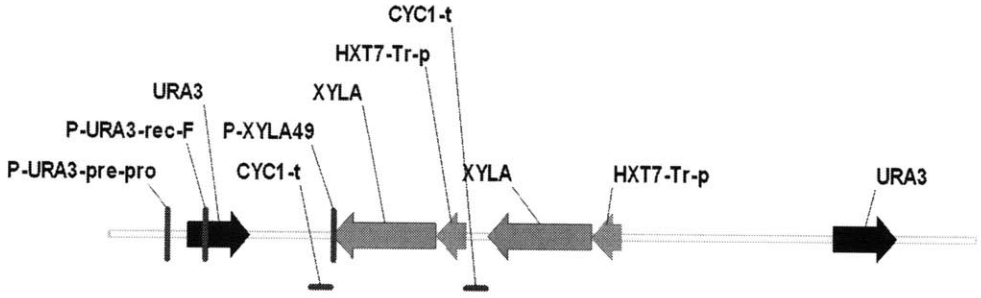


Figure 5-6. $(XYLA)_2$ integration in *URA3* of H142-p406A2

In H142E-p406A2, the primer P-XYL49 and two sense primers P-URA3-pre-pro (5'- GGTTAATGTGGCTGTGGTTTCAGGGTCC) and P-URA3-rec-F (5'- CCGCTAAAGG CATTATCCGC) were used to amplify either a 2.2 or a 1.7 kbps fragment. However, only a PCR product sized about 1 kbps could be achieved using primer P-XYL49 and P-URA3-pre-pro. The sequencing result showed a structure of truncated *URA3* flanked by a retrotransposon gene Ty1 sequence (Figure 5-7), suggesting that some genome rearrangement occurred during the evolution. The structure and location of *XYLA* duplication will be investigated in the following sections by Southern and chromosome blots.

Figure 5-7. Sequencing of *URA3* context in H142-p406A2

A quantitative PCR was used to evaluate the copy numbers of *XYLA* in the H142E strains, as shown in Figure 5-8.

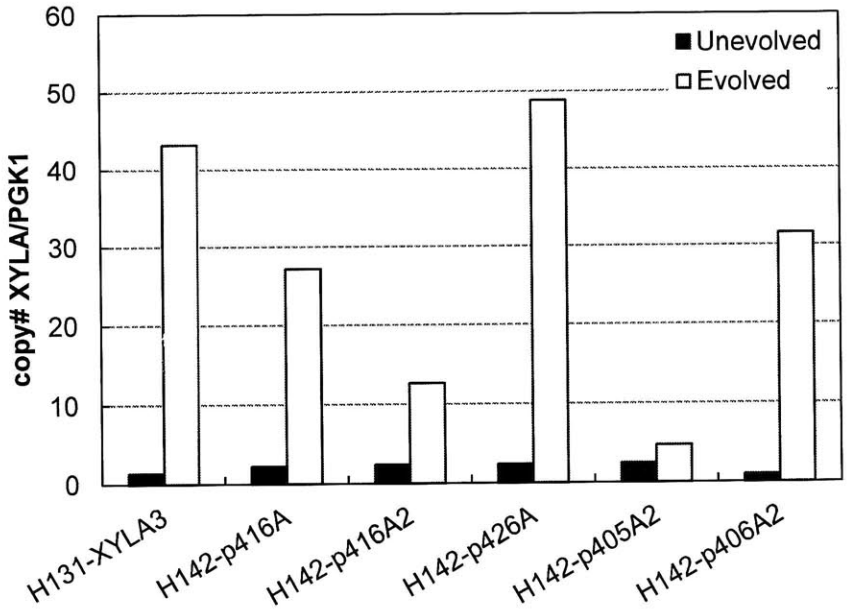


Figure 5-8. qPCR for *XYLA* copy numbers in H142/H142E strains

As expected, elevated copy numbers (12-49) of *XYLA* in all of the H142E strains were revealed, as compared with 0.9-2.4 copies in the unevolved H142 strains.

5.3.2 Southern blot

The Southern blot was run for the evolutionarily engineered H131E and the newly evolved H142E strains as well as unevolved H131 and H142 as references.

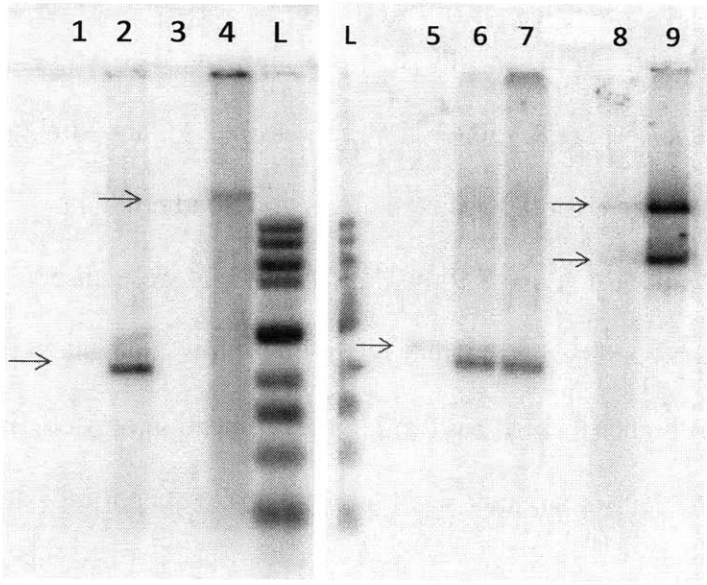


Figure 5-9. Southern blot of *XYLA* in H131/H142 strains

The chromosomal DNA of evolved H131E5-XYLA31 and its reference H131-XYLA31 were digested by *Xba* I and *EcoR* I before the blot (Figure 5-9, Lanes 1-4). For both pairs, a higher intensity (7-12.5× over reference) of the *XYLA* blot was observed. Moreover, the 2.2k bps band (Lane 2) of *Xba* I digested H131E5-XYLA31 corresponds to the size of the *TDH3p-XYLA-CYC1t* construct, which indirectly verifies the tandem gene duplication of *XYLA* in the evolutionarily engineered strain, as shown in Figure 5-10.

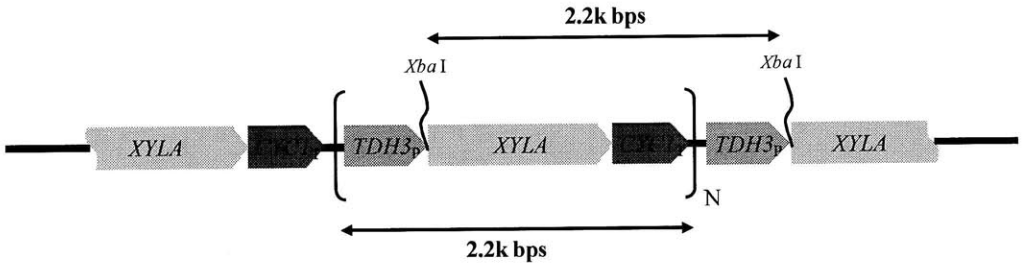


Figure 5-10. Structure of tandem *XYLA* duplication after *Xba* I digestions

Xba I digestion was also applied on other H131E and H142E strains. As Lane 5-7 shows, the *XYLA* blot intensities of H131E5 and H131E8 are very similar, about 12× the reference. This observation is consistent with the qPCR results showing that the copy numbers of *XYLA* in H131E5 and H131E8 (50.6±14.4 and 43.7±1.1) are similar. Nevertheless, there is a 4-fold systematic error between the qPCR and Southern blot.

For the new evolved H142E-p416A (Lane 8-9), two blots of *XYLA* were observed on the evolved strain and its reference. The I_9 (the sum of the two blots' intensities in Lane 9) is approximately 9.2 times higher than I_8, confirming the duplication of *XYLA* in H142E-p416A. However, the two blots of various sizes (other than 2.2 kbps) indicate different structures than the tandem *XYLA* duplication in the H131E-XYLA31 strains.

5.3.3 Transcription of key enzymes in the xylose metabolic pathway

The transcription levels of the key xylose metabolic genes were determined through quantitative RT-PCR, for the evolutionarily engineered strains H131E5-XYLA31 and H131E8-XYLA31 and control strain H131-XYLA31. The constitutively expressed endogenous *PGK1* was used as a reference.

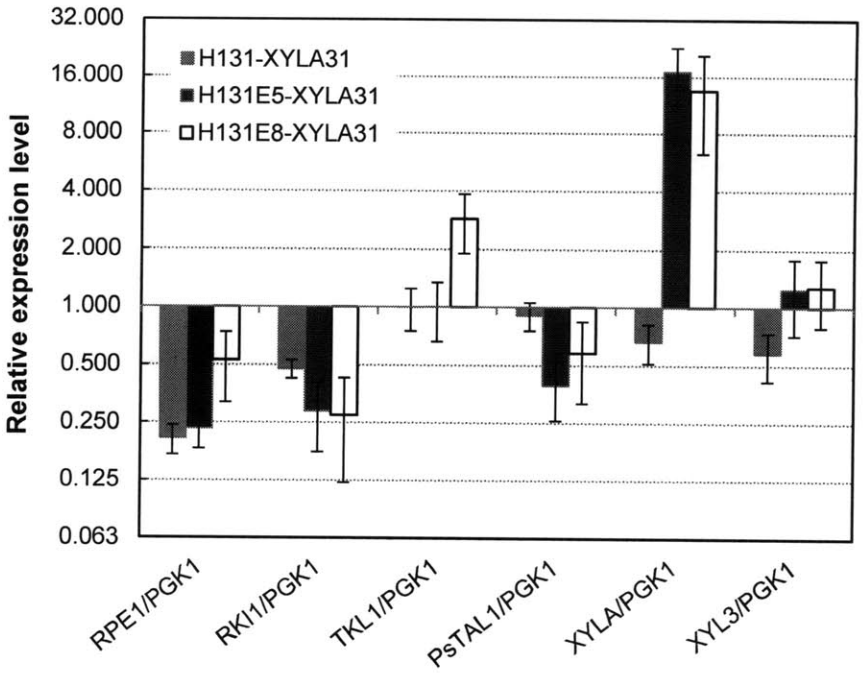
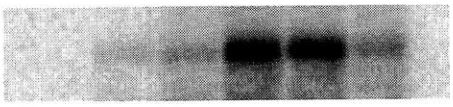


Figure 5-11. Transcription of the key xylose metabolic genes

The expression profiles of H131E5-XYLA31 and H131E8-XYLA31 are very similar, and the most significant change for both evolved strains in comparison to the reference is the high *XYLA* transcription level. The transcription of *XYLA* in H131E5/E8 are 17.1±5.8 and 13.6±7.3 times higher than that of *PGK1*, or 25.7±8.8 and 20.4±11.0 times higher than that of *XYLA* in unevolved H131-XYLA31, while the *XYL3* expression level of H131E5/E8 is only 2-fold higher than that of the control strain H131-XYLA31. The only difference between H131E8 and H131E5 is the 2-fold higher transcription of *RPE1* and *TKL1* in H131E8.

Northern blot, another way to evaluate *XYLA* transcription, was run for the series of H131E evolutionarily engineered strains in comparison to their parent strain H131, as shown in Figure 5-12.



0: H131
1: H131-XYLA31
2: H131E1-XYLA31 ($I_2 = 2.1 \times I_1$)
3: H131E3-XYLA31 ($I_3 = 7.9 \times I_1$)
4: H131E5-XYLA31 ($I_4 = 6.8 \times I_1$)
5: H131E8-XYLA31 ($I_5 = 3.0 \times I_1$)

I_X: intensity of band indicated in lane x

Figure 5-12. Northern blot results of *XYLA* in H131E strains

All the H131E strains showed improved *XYLA* transcription. Nevertheless, the transcription level did not monotonically increase along the evolution, suggesting that the elevated expression level of *XYLA* played a dominant role only during the earlier phase of the evolution.

5.3.4 Enzymatic activity of *XYLA* and *XYL3*

In vitro XI activities of two H131E-XYLA31 strains were also compared with the parent strain of H131-XYLA31. Consistent with the DNA copy number and transcription level of *XYLA*, XI activity had significantly increased in the evolved isolates H131E5-XYLA31 and H131E8-XYLA31, about 13-fold higher than that of the unevolved strain H131-XYLA31 (Figure 5-13). No activity could be detected in strain H131 without *XYLA*.

A similar comparison was done on XK, as shown in Figure 5-13. Very weak XK activity (0.022 U/mg) could be detected in strain H131, due to the expression of the endogenous *XKS1*. After overexpression of *PsXYL3*, the XK activity of H131-XYLA31 was increased to 1.41±0.07 U/mg. In contrast to the elevated XI activity after evolutionary engineering, the XK activity decreased to 0.5-0.9 U/mg for H131E5/E8-XYLA31, resulting in a much more balanced XI/XK activity ratio.

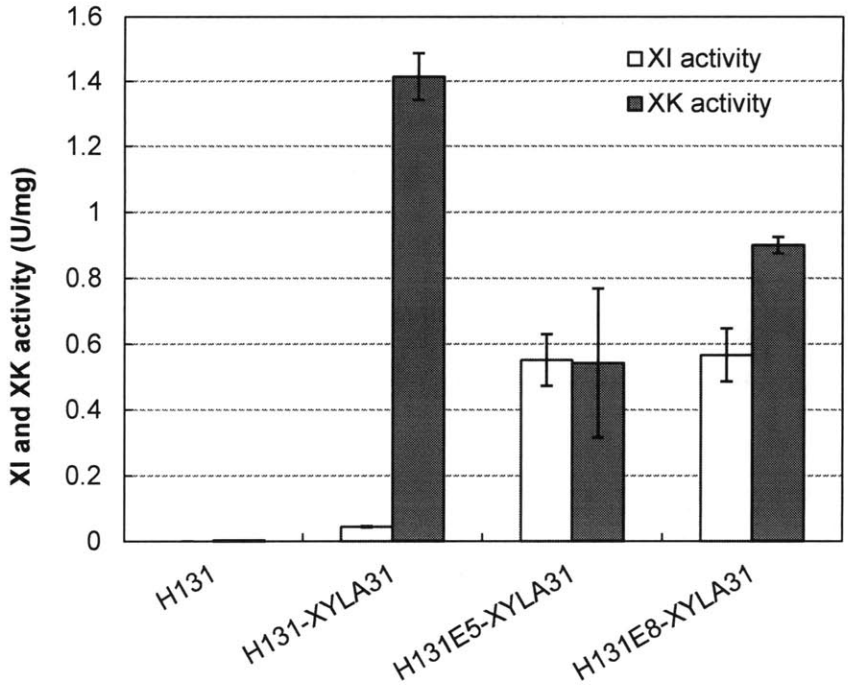


Figure 5-13. Activity of XI and XK in H131/H131E strains

All the molecular and biochemical characterizations of *XYLA* and *XYL3* (normalized to reference strain H131-XYLA31) are summarized in Figure 5-14.

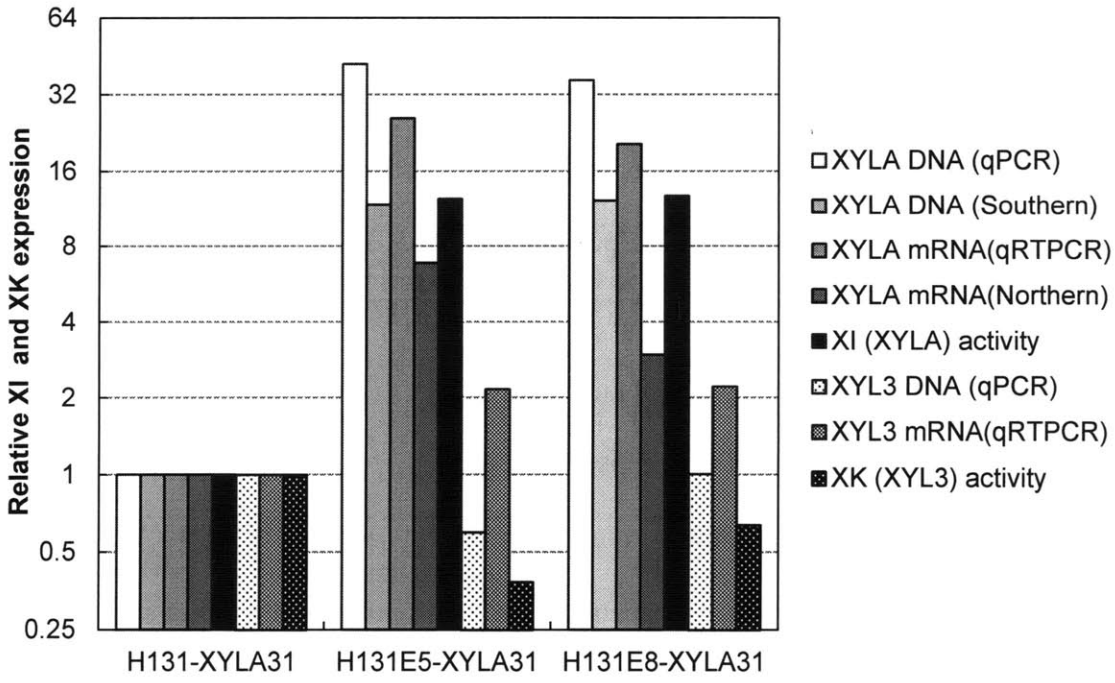


Figure 5-14. DNA copy number, expression and activity of XI and XK in H131/H131E strains

For *XYLA*, the gene copy number and expression level of both evolved strains are 8-32 times higher than the reference strain, with a good correlation of DNA, RNA and protein activity. The only exception is the Northern blot of *XYLA* in H131E8-XYLA31, which could be due to inaccuracy in the experiment and lack of replication. For *XYL3* the DNA, RNA, and protein levels of the evolved strains are comparable to the reference strain with good correlation (a system error of about 2-fold).

5.3.5 Chromosome blotting

Tandem *XYLA* duplication was confirmed in the evolved H131E strains by use of the Southern blot and qPCR. In addition, failure to segregate plasmid-carrying *XYLA* after evolutionary engineering implies that the gene may have been recombined to chromosomal DNA. To determine the location of *XYLA*, pulsed field gel electrophoresis (PFGE) was run for H131E-XYLA31 and the new evolved H142E-XYLA strains, as shown in Figure 5-15.

In the unevolved H131-XYLA31, a very faint blot could be seen at the position of Chr III. In H131E1-XYLA31, a strong blot on Chr X suggests multiple copies of *XYLA* had been recombined to the chromosome. More interestingly, H131E3/E5/E8-XYLA31 displayed a similar blot pattern on the position of Chr VI with increased size. The new size of Chr VI is between Chr III (317 kbps) and Chr IX (440 kbps), more than 50 kbps larger than the original size of 270 kbps. The difference of the size is consistent with the size of multiple copies of *XYLA* (2.2 kbps×30-50 copies).

The different *XYLA* integration locations of H131E3/E5/E8-XYLA31 and H131E1-XYLA31 indicate either independent evolution process of two branches or some translocation of *XYLA* between H131E1-XYLA31 and H131E3-XYLA31.

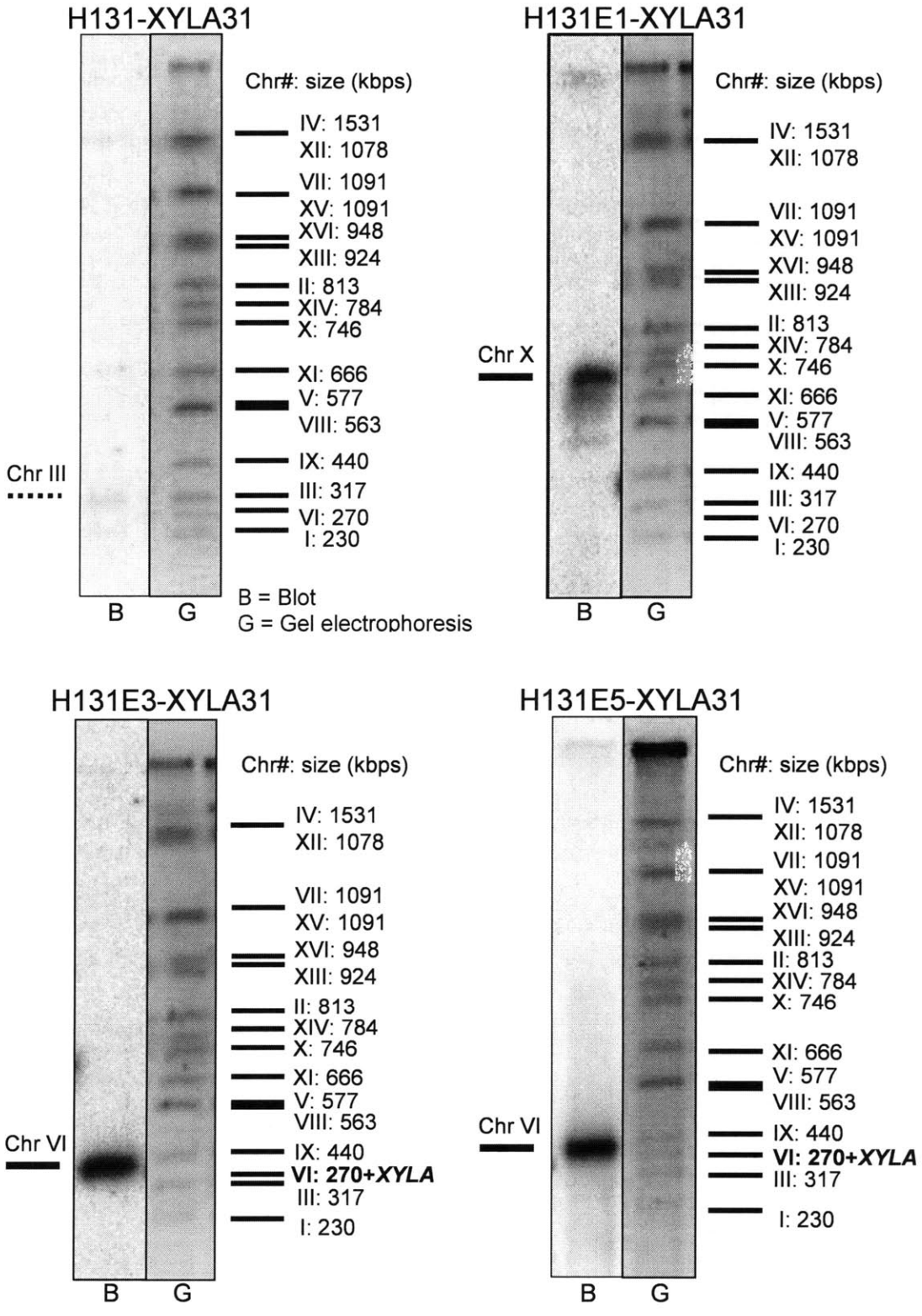


Figure 5-15. PFGE and chromosome blot of H131/H131E/H142E

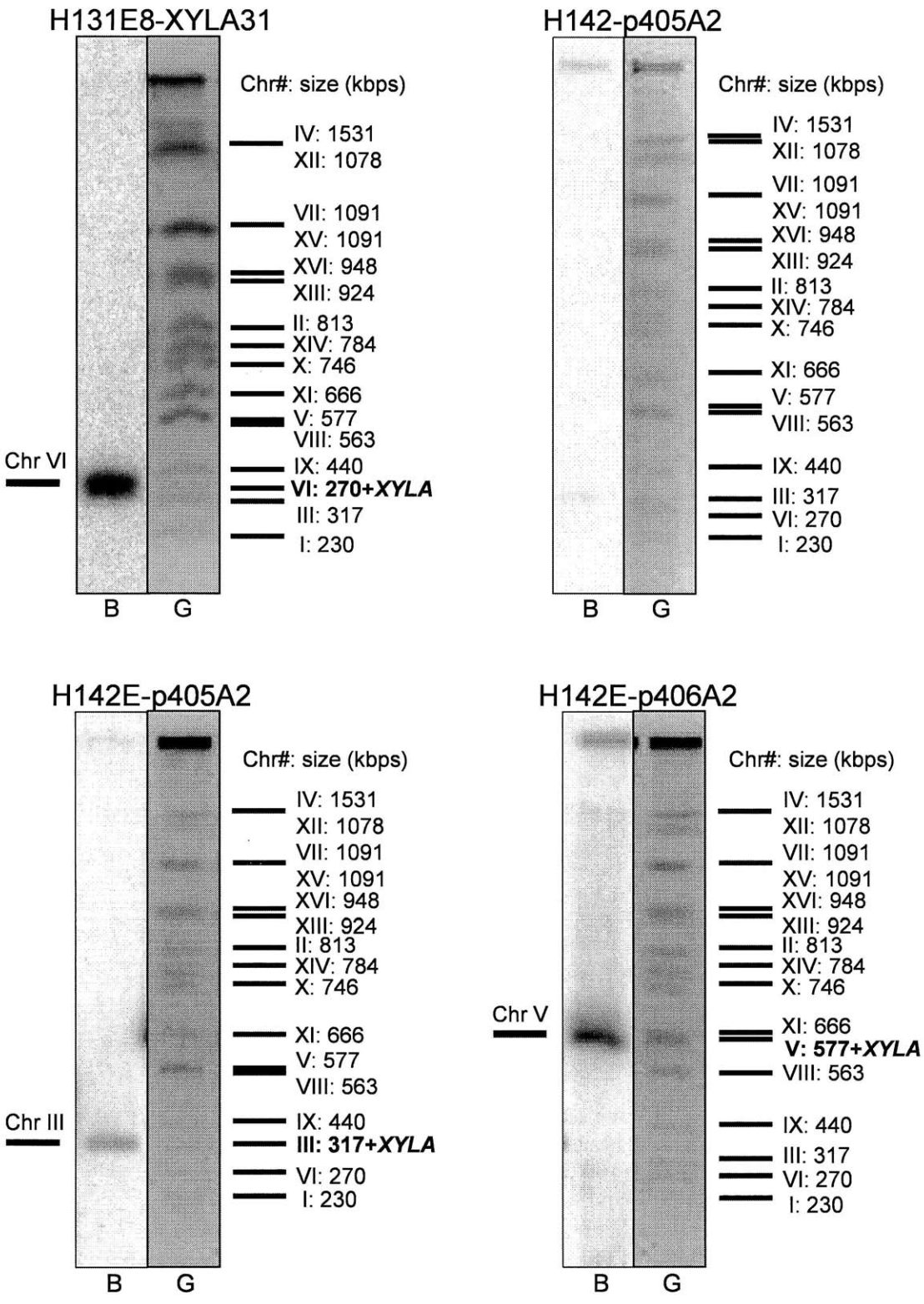


Figure 5-15. PFGE and chromosome blot of H131/H131E/H142E (continued)

The H142E-p405A2 and H142E-p406A2 were confirmed to have *XYLA* integrated on the LEU2 (Chr III) and URA3 (Chr V) loci as expected, while the *XYLA* blot could hardly be detected in the unevolved H142-p405A2 and H142-p406A2 (not shown).

For H142E-p405A2, the *XYLA* duplication structure was detected to be in Chr III on the locus of *LEU2*, which was confirmed by PCR previously. The Chr III size migrated slightly but obviously, in agreement with the minor *XYLA* copy number increment determined by qPCR.

For H142E-p406A2 the *XYLA* duplication structure was detected in the Chr V where *URA3* locates, although the context of *XYLA* integration could not be verified by using PCR. The Chr V size migration was about 70k bps (577k bps to ~650k bps), in agreement with the copy number determined by qPCR (2.2k bps/copy × 32 copies ≈ 70k bps).

Although the presence of tandem *XYLA* duplication was not verified directly, the evidence of chromosome blot, qPCR results, and the tandem gene duplication structure on pRS415-geno-W2 strongly suggested that the duplication and integration of gene *XYLA* was a major recombination event during the evolutionary engineering, and a high expression level of heterologous XI is necessary for rapid xylose utilization.

5.3.6 *GRE3* mutations in the evolved strains

The XI activity is inhibited by xylitol [65]. Although natural *S. cerevisiae* strains do not have a specific xylose reductase (*XYL1*), xylose could be reduced to xylitol with an endogenous aldose reductase like *GRE3* [63] and other putative XR [64, 114] listed in Table 5-2.

Table 5-2. Putative xylose reductases in *S. cerevisiae*

Gene	Description
GRE3	Aldose reductase involved in methylglyoxal, d-xylose, arabinose, and galactose metabolism; stress induced (osmotic, ionic, oxidative, heat shock, starvation and heavy metals); regulated by the HOG pathway
YPR1	NADPH-dependent aldo-keto reductase
GCY1	Putative $NADP^+$ coupled glycerol dehydrogenase
YJR096W	Putative xylose and arabinose reductase
YDL124W	NADPH-dependent alpha-keto amide reductase

Deletion of the *GRE3* gene improves efficient xylose utilization in XI-expressing *S. cerevisiae* strains [66] and also reduces xylitol yields [20, 67]. However, it has been reported that the aldose reductase encoded by the *GRE3* gene belongs to a group of generally stress-induced proteins [63] and that its deletion reduces the growth by 30% [67]. Based on the above analysis and estimation of the xylitol effect shown in the Appendix, the *GRE3* was not knocked out during the strain construction.

The *GRE3* in the evolved strains were sequenced in order to evaluate the inhibitory effect of xylitol (*GRE3* product) on xylose utilization. Interestingly, mutations were found to either truncate the C-terminus of the *GRE3* protein by 28 amino acids (H131E1-XYLA31) or switch Glu193 to Lys (H131E3/E5/E8-XYLA31), as shown in Figure 5-16.

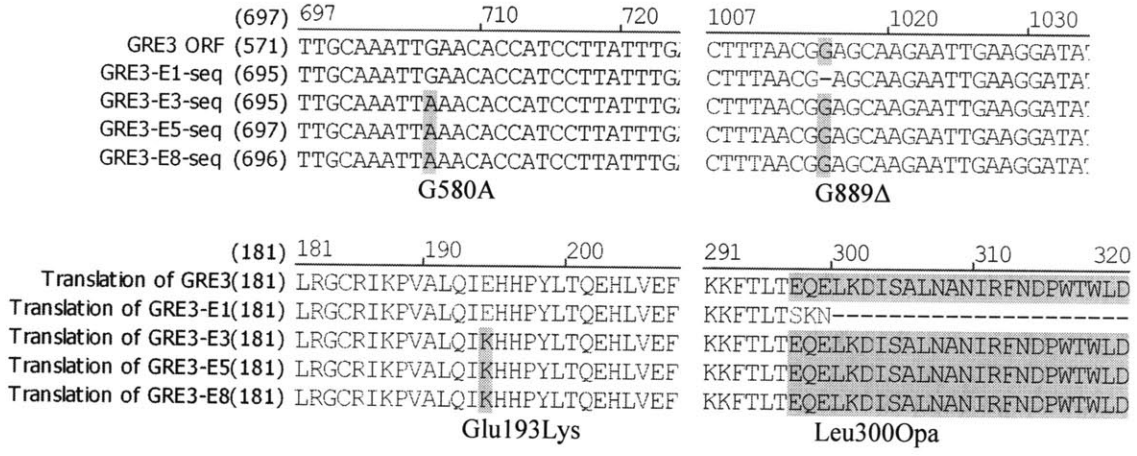


Figure 5-16. Mutations of *GRE3* in the evolved strains

It can be inferred that the mutations of *GRE3* lessen its xylose reductase activity and therefore reduce the xylitol accumulation, which is favorable for xylose assimilation since it is a competitive inhibitor of xylose isomerase.

In addition, the different *gre3* mutations in H131E1-XYLA31 and H131E3/E5/E8-XYLA31 coincide with the different *XYLA* integration locations (Figure 5-15) of the two sets of strains, implying that H131E1-XYLA31 and H131E3/E5/E8-XYLA31 evolved independently in two branches.

Furthermore, the following experiments could be run to further evaluate the *GRE3* effects: i) the XR activity of $gre3^m$(E193K) and $gre3^m$(L300Opal); ii) xylitol inhibitory kinetics in the unevolved and evolved strains; and iii) comparison of *GRE3*, *gre3Δ* and $gre3^m$ in terms of XR activity, xylitol accumulation, and growth rate.

To first validate the effect of *gre3Δ* on growth rate, two strains (H131-LEU2 and H131-gre3Δ, Table 5-1) were constructed. In the H131-gre3Δ a copy of *LEU2* was inserted into the *GRE3* ORF at 522, and it truncated the *GRE3* by about half of the amino acid sequence. H131-LEU2 was constructed as a control that has the same genetic background (except for *LEU2* on its original locus other than in the *GRE3*). The growth rates of the two strains were compared on SDG medium under micro-aerobic conditions, as shown in Figure 5-17.

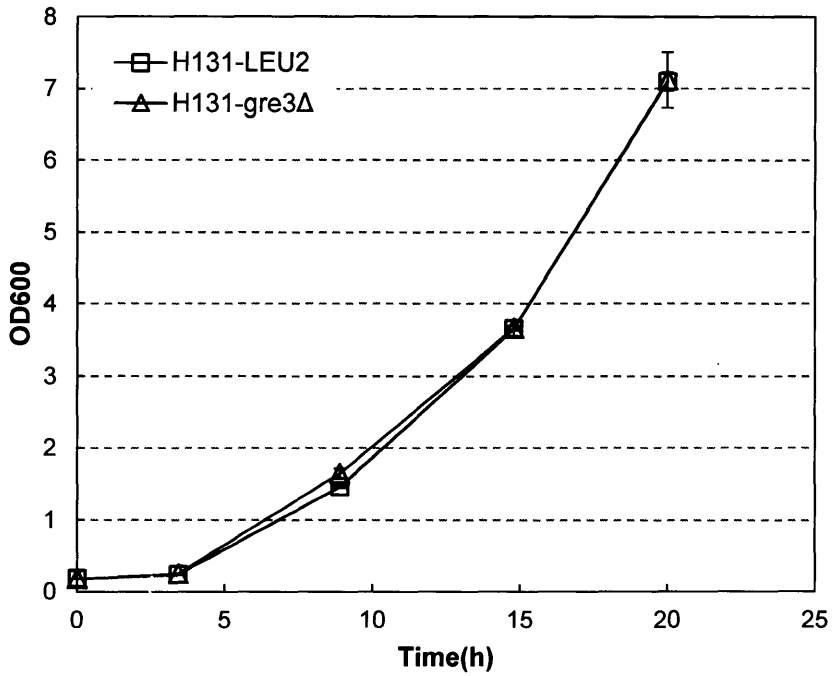


Figure 5-17. Effect of *gre3Δ* on aerobic growth in glucose medium

The comparison of H131-gre3Δ and H131-LEU2 showed no difference in growth, unlike findings in a previous report [67]. Similar results were confirmed by another set of strains on glucose (data not shown). However, to ultimately verify the effect of *GRE3*, a set of strains containing *GRE3*, *gre3Δ*, and $gre3^m$ need to be constructed based on the evolved xylose-fermenting strain and evaluated on xylose.

5.4 Conclusions

A high expression level of XI as a key phenotype was confirmed by quantitative PCR, blotting and hybridization techniques, and enzyme activity assay. The tandem gene duplication (TGD) of *XYLA* was proved in the series of evolved H131E-XYLA31 strains. The TGD structure was translocated onto one or more chromosomes during the early stage of the evolutionary engineering. However, the TGD is not the only way to achieve high XI expression. The evolution process was successfully reproduced in a different background H142, and *XYLA* expansion could originate from a different initial *XYLA*

construction, whether by integration or on different types of plasmids. The location-specific *XYLA* integration and duplication could be achieved, although the structure might be subject to other genomic rearrangements.

The XK (*XYL3*) form *P. stipitis* could be efficiently expressed in *S. cerevisiae*, and the low copy number of *PsXYL3* promoted by a strong promoter should be enough to support the high flux of xylose assimilation. Mutations on *GRE3* that potentially abolish XR activity were discovered in the evolutionary engineered strains, supporting the hypothesis that the elimination of XR activity was favorable for xylose fermentation by reducing the production of by-product xylitol and hence alleviating inhibition to XI.

CHAPTER 6. Optimization of Fermentation and Other Metabolic Engineering of Xylose-fermenting *S. cerevisiae* Strains

6.1 Introduction

Successful evolutionary engineering of *S. cerevisiae* resulted in strain of H131E8-XYLA31, which showed comparable fermentation performance to the best-reported strain, RWB-218 [3, 21], under a typical yeast cultivation conditions. Since the cultivation conditions may not be optimal for the anaerobic xylose fermentation, optimization of the conditions was expected to improve the fermentation performance of the evolutionarily engineered strain. Moreover, as one of the best xylose-utilizing strains, H131E8-XYLA31 could be used as the host for further rational metabolic engineering.

In this chapter, fermentation condition including media, strictly anaerobic setup and chemostat cultivation were studied to further explore the best performance of the engineered strain. After restoring the auxotrophic markers *arg4* and *leu2* (primarily *leu2*) and supplementing the anaerobic growth factors (ergosterol and Tween 80) to the media, the newly engineered strain H153E10-XYLA31 showed a greatly improved anaerobic growth rate of 0.199 h^{-1} and specific xylose consumption rate of 1.647 $g \cdot g^{-1}h^{-1}$, 65% and 37% higher than those of the best reported xylose-fermenting strain RWB 218, respectively. In addition, several genes that improve xylose utilization or cell growth were introduced to the engineered strain, including an NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (*gapN*), some other heterogeneous xylose isomerase and a series of putative xylose transporters.

6.2 Materials and Methods

6.2.1 Plasmids and strains

Plasmids used in this chapter are summarized in Table 6-1 and yeast strains constructed in this chapter are summarized in Table 6-2.

Table 6-1. Plasmids used in this chapter

Plasmids and strains	Characteristics
pUCAR1	pUC19, *Hind*III-*ARG4-Hind*III
pUCAR1-gapN	pUCAR1, *BamH* I-*TDH3p-gapN-CYC t -EcoR* I
pRS405	YIp, *LEU2*
pRS405-$(XYLA)_2$	*LEU2,* 2 copies of $HXT7_{Tr}$*p-XYLA-CYC1t* in tandem
pRS406-$(XYLA)_2$	*URA3,* 2 copies of $HXT7_{Tr}$*p-XYLA-CYC1t* in tandem
Heterologous XIs	
pRS426HXT7$_{Tr}$-PXYLA	2μ ori, *URA3, Piromyces XYLA*
pRS426HXT7$_{Tr}$-TTHXYLA	2μ ori, *URA3, Thermus thermophilus XYLA*
pRS426HXT7$_{Tr}$-EcXYLA	2μ ori, *URA3, E.coli XYLA*
pRS426HXT7$_{Tr}$-BTHXYLA	2μ ori, *URA3, Bacteroides thetaiotaomicron XYLA*
pRS426HXT7$_{Tr}$-AthXYLA	2μ ori, *URA3, Arabidopsis thaliana XYLA*
pRS426HXT7$_{Tr}$-BsXYLA	2μ ori, *URA3, B.subtilis XYLA*
pRS426HXT7$_{Tr}$-SrXYLA	2μ ori, *URA3, Streptomyces rubiginosus XYLA*
pRS426HXT7$_{Tr}$-XcaXYLA	2μ ori, *URA3, Xanthomonas campestris XylA2*
Putative xylose transporters	
p424-TEF-DhSUT2	*Debaryomyces hansenii SUT2*
p424-TEF-DhSUT3	*Debaryomyces hansenii SUT3*
p424-TEF-CaSUT2	*Candida albicans SUT2*
p424-TEF-PsXUT4	*Pichia stipitis XUT4*
p424-TEF-PsSUT1	*Pichia stipitis SUT1*
p424-TEF-PsSUT2	*Pichia stipitis SUT2*
p424-TEF-DhSUT1	*Debaryomyces hansenii SUT1*
p424-TEF-CaSUT1	*Candida albicans SUT1*
p424-TEF-CgSUT1	*Candida glabrata, SUT1*
p424-TEF-PsXUT6	*Pichia stipitis XUT6*
p424-TEF-PsXUT5	*Pichia stipitis XUT5*
p424-TEF-TrXLT1	*Trichoderma reesei XLT1*

Table 6-2. *S. cerevisiae* strains used in this chapter

S. cerevisiae Strains	Characteristics
H131E3-XYLA31	H131E-XYLA31, isolated from micro-aerobic sequential batch cultivation
H131E10-XYLA31	Similar to H131E8-XYLA31, isolated from anaerobic chemostat 2 weeks later than H131E8-XYLA31
H145E10-XYLA31	H131E10-XYLA31, *arg4* complemented by pUCAR1 containing *ARG4*
H146E10-XYLA31	H131E10-XYLA31, *leu2-3*complemented by pRS405 containing *LEU2*
H153E10-XYLA31	H146E10-XYLA31, *arg4* complemented by pUCAR1 containing *ARG4*
H131E5-XYLA31-ARG4	H131E5-XYLA31, *arg4* complemented by pUCAR1 containing *ARG4*
H141E5-XYLA31	H131E5-XYLA31, pUCAR1-gapN
H142	H131, *ARG4::HXT7p-PsXYL3-CYC1t*
H142-XYLAs	H142, pRS426HXT7$_{Tr}$-XYLAs, XYLAs listed in Table 6-1
H142E-XYLAs	H142-XYLAs, evolutionary engineered by aerobic sequential batch cultivation
FPL-YSX3C[17]	MAT_α, *trp1-112*, *leu2 ::LEU2-TDH3p-XYL1*, *ura3::URA3- TDH3p - XYL2*, Ty3::*NEO-XYL3*
YSX3-SUTs	FPL-YSX3C, with putative xylose transporters listed in Table 6-1

6.2.2 Anaerobic fermentation of xylose

The condition of anaerobic fermentation is described in Chapter 3. In addition, to achieve strict anaerobic condition, cultures were stirred at 300 rpm and sparged with 0.5 L/min nitrogen (Airgas, Ultra high purity grade, <1ppm oxygen). And the medium was supplemented with 200 μL/L of silicone antifoam as well as with the anaerobic growth factors ergosterol (0.01 g/L) and Tween 80 (0.42 g/L) dissolved in ethanol in 1000×.

6.3 Results and Discussion

6.3.1 Medium comparison

Yeast nitrogen base (YNB) is used as a base medium for preparation of minimal and synthetic defined yeast media, which contains basically salts, vitamins and trace elements. To evaluate the effect of YNB and other supplemental media, a series of media were tested using H131E3-XYLA31 under micro-aerobic condition and 2% xylose unless

otherwise indicated. The growth and xylose consumption after 24 hours were compared in Figure 6-1.

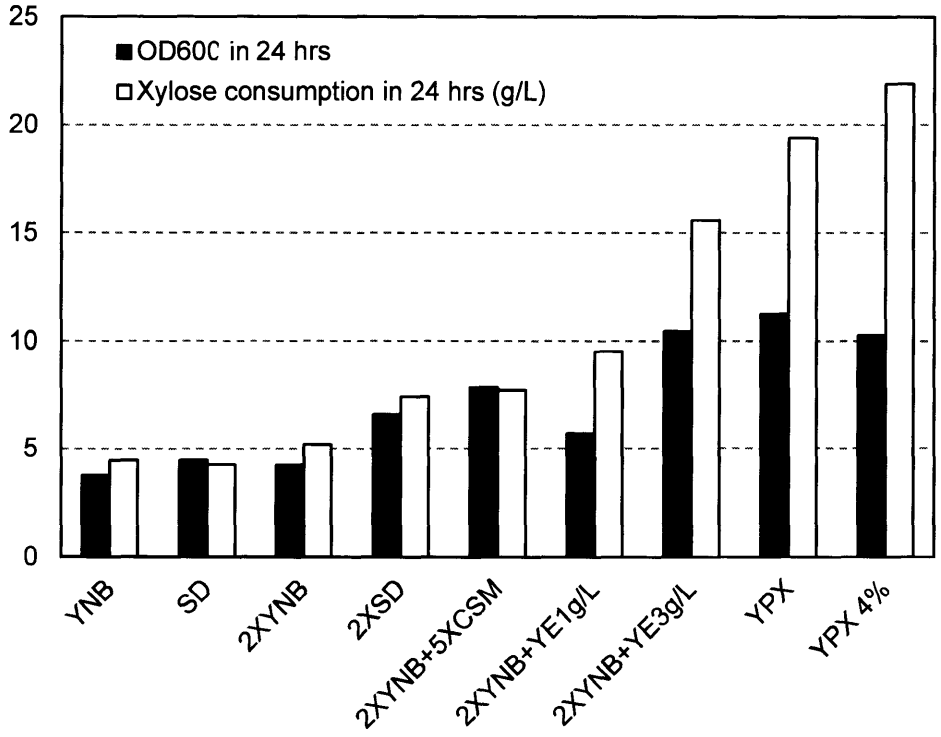


Figure 6-1. Fermentation of H131E3-XYLA31 on different media

Additional YNB did not appreciably contribute to the growth. On top of YNB, a supplement of 1-5× CSM (mixture of essential amino acids and nucleobases) improved the growth and xylose consumption by about 2-fold, implying that amino acid synthesis under oxygen-limited condition could be one limiting factor. Furthermore, 1-10 g/L YE (yeast extract) greatly improves growth by up to 3-fold and hence xylose utilization by up to 5-fold, suggesting that some growth factor was deficient in the minimal YNB media and could be supplemented by the addition of complex YE, which contains most essential nutrients for yeast growth.

6.3.2 Complement of *arg4* and *leu2*

The evolutionarily engineered strain H131E10-XYLA31 still contains two auxotrophic markers, *arg4* and *leu2*, and requires L-arginine and L-leucine in the medium. To evaluate the performance of the strain on the minimal medium without Arg and Leu, the two genes were complemented by YIp pRS405 and pUCAR1 (Table 6-1), separately.

H131E10-XYLA31 was first transformed with *Cla* I linearized pUCAR1 and selected on Arg$^-$ medium to yield H145E10-XYLA31. Meanwhile, H131E10-XYLA31 was also transformed with *Afl* II digested pRS405 and selected on Leu$^-$ medium to yield H146E10-XYLA31. H146E10-XYLA31 was then transformed with *Cla* I linearized pUCAR1 to yield H153E10-XYLA31 (Table 6-2). The aerobic growth profiles of the three new constructed strains were compared with their parent strain H131E10-XYLA31 on 2×YNBX medium (with 2× arginine and/or leucine if necessary) to verify the effect of *ARG4* and *LEU2* on growth, as shown in Figure 6-2.

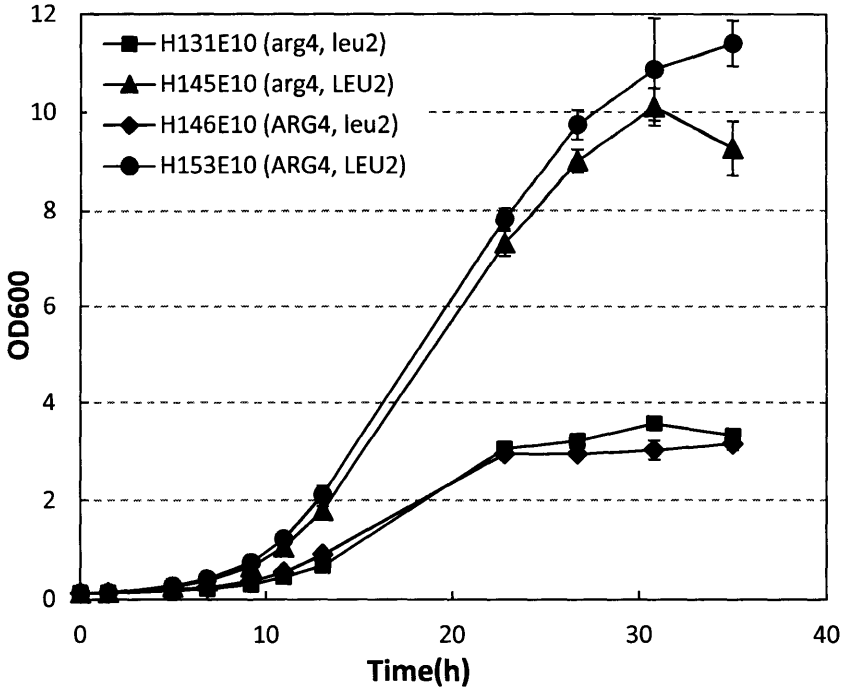


Figure 6-2. Effect of *ARG4* and *LEU2* on aerobic growth

The successful restoration of the *ARG4* gene eliminates the requirement of arginine supplement in the medium but didn't change the growth profile significantly, when comparing the performances of H131E10 with H146E10, or H145E10 with H153E10.

Surprisingly, the restoration of *LEU2* not only fixed the deficient in the leucine biosynthesis, but also increased the μ_{max} by 30% and elevated the final biomass by 3-fold, when comparing the strains H145E/H153E (*LEU2*) with H131E/H146E (*leu2*). It had been reported that the amount of leucine provided in commonly used synthetic media (like the YNB used in this study) is limiting for growth of leucine-requiring strains, and supplementing synthetic media with a higher leucine concentration is recommend [115]. Nevertheless, the leucine concentration still limited the growth even though it had been increased to 200mg/L, 2× the concentration of normal SD media. In addition, studies have demonstrated that leucine was also correlative with tolerance to ethanol [116]. Therefore, a complement of *LEU2* solves both problems and dramatically boosts growth on xylose.

The performance of the prototrophic H153E10-XYLA31 in terms of anaerobic fermentation was evaluated in a bioreactor in 2×YNBX (4%) medium, to compare with that of H131E8-XYLA31 under similar condition (fermentation results from Chapter 3), as shown in Figure 6-3 and Table 6-3.

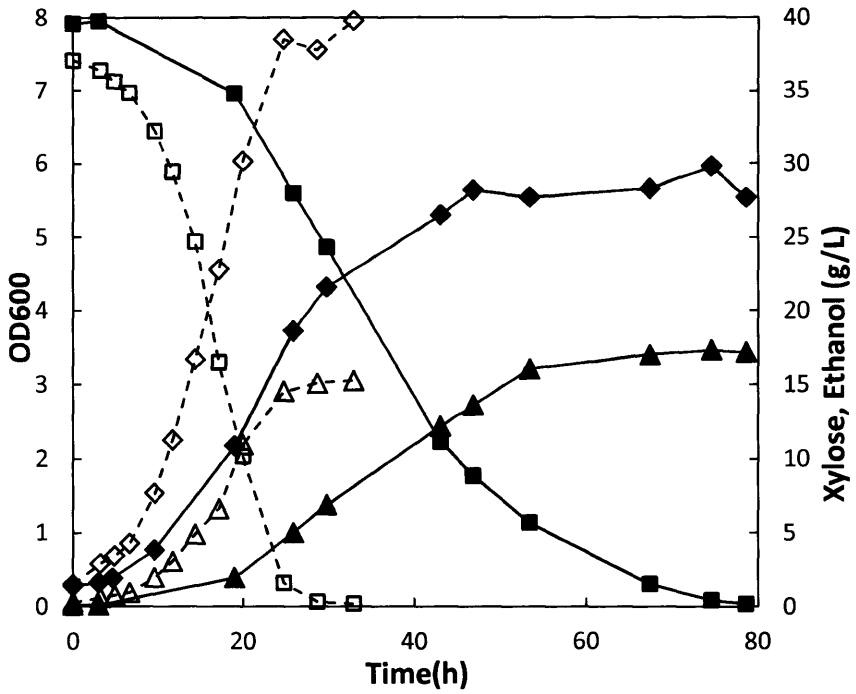


Figure 6-3. Comparison of H131E8-XYLA31 and H153E10-XYLA31

H131E8-XYLA31: (◇OD_{600}, □Xylose, △Ethanol)

H153E10-XYLA31: (◆OD_{600}, ■Xylose, ▲Ethanol)

Table 6-3. Characteristics of H131E8-XYLA31 and H153E10-XYLA31 fermentation

		H131E8-XYLA31		H153E10-XYLA31	
		Growth phase	Overall	Growth phase	Overall
μ_{max}	h^{-1}	0.120±0.004	\	0.177±0.009	\
Δt	h	40.05	71.60	19.85	29.70
Δ[Xylose]	g/L	28.62	39.35	34.02	36.19
Δ[Ethanol]	g/L	12.10	17.22	13.75	14.69
Δ[DCW]	g/L	1.327	1.503	1.869	1.968
Δ[Glycerol]	g/L	1.057	1.565	2.114	2.134
Final OD_{600}		\	5.540	\	7.960
Average OD600		2.926	4.117	3.780	4.675
$DCW_{average}$	g/L	0.78	1.098	1.008	1.336
q_{xylose}	$g \cdot L^{-1} h^{-1}$	0.715	0.550	1.714	1.219
	$g \cdot g^{-1} h^{-1}$	0.916	0.501	1.700	0.912
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	0.302	0.240	0.693	0.495
	$g \cdot g^{-1} h^{-1}$	0.387	0.219	0.687	0.370
Ethanol yield	g/g	0.423	0.438	0.404	0.406
Biomass yield	g/g	0.046	0.038	0.055	0.054
Glycerol yield	g/g	0.037	0.040	0.062	0.059

The newly constructed strain H153E10-XYLA31 finished the 37 g/L xylose in about 37 hours, twice as fast as H131E8-XYLA31 (40 g/L xylose in 80 hours), thanks to the 47.5% faster growth, 43% higher final biomass and 120% higher specific xylose consumption rate. The xylose was almost depleted when the growth stopped, and likely limited further biomass accumulation. Therefore, better biomass accumulation could be achieved if sufficient carbon source xylose is provided.

Furthermore, H153E10-XYLA31 has 42% higher biomass yield and 8% lower ethanol yield. It was suspected that oxygen leaked into the system and led to better growth and lower fermentation efficiency.

To address the above issues, batch fermentation was carried out with a higher initial xylose concentration of 80g/L and carefully insulated from oxygen diffusion. The results are shown in Figure 6-4 and Table 6-4.

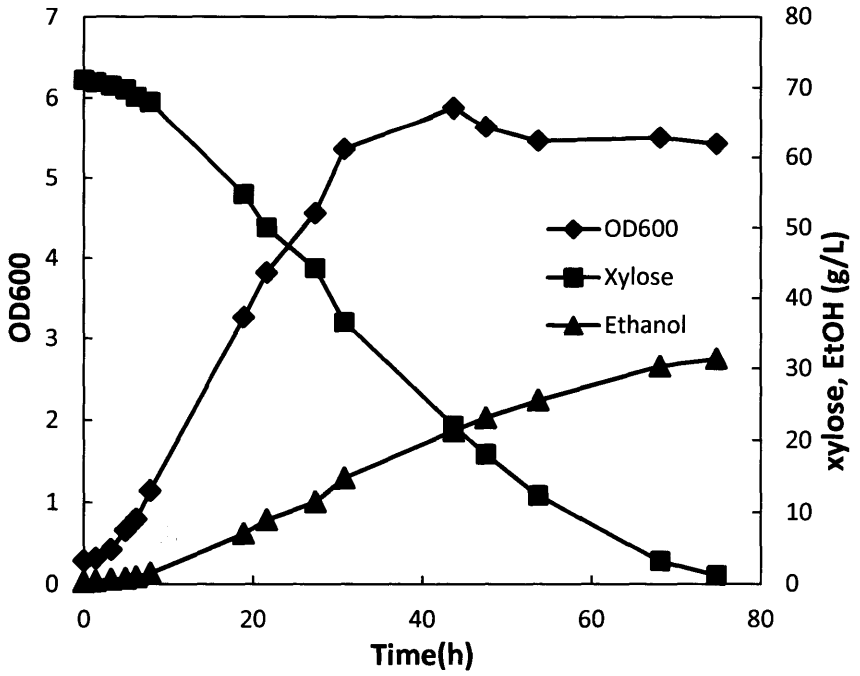


Figure 6-4. Fermentation of H153E10-XYLA31 with 8% xylose

Table 6-4. Characteristics of H153E10-XYLA31 fermentation with 8% xylose

		Growth phase	Overall
μ_{max}	h^{-1}	0.202±0.009	/
Δt	h	24.15	74.80
Δ[Xylose]	g/L	26.08	69.94
Δ[Ethanol]	g/L	10.83	30.99
Δ[DCW]	g/L	1.103	1.369
Δ[Glycerol]	g/L	1.530	2.799
$OD_{average}$		2.547	4.332
$DCW_{average}$	g/L	0.679	1.155
q_{xylose}	$g \cdot L^{-1} h^{-1}$	1.080	0.935
	$g \cdot g^{-1} h^{-1}$	1.590	0.809
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	0.448	0.414
	$g \cdot g^{-1} h^{-1}$	0.660	0.359
Ethanol yield	g/g	0.415	0.443
Biomass yield	g/g	0.042	0.020
Glycerol yield	g/g	0.059	0.040

Under anaerobic conditions, H153E10-XYLA31 rapidly finished 71 g/L xylose within 75 hours at overall ethanol yield of 0.443. The initial growth (μ_{max} = 0.202±0.009 h^{-1}) and the specific xylose productivity in growth phase (1.590 $g \cdot g^{-1} h^{-1}$) were also very good. Nonetheless, the biomass accumulation stopped at OD_{600} about 6 when the residual xylose of more than 20 g/L should not limit the growth. The final cell density is only about a half of that of the same strain on the same medium under aerobic condition (H153E10-XYLA31 in Figure 6-2). The restrained growth could be due either to a deficiency of some growth factor under the oxygen-limited condition or to some stress caused by accumulation of certain metabolites.

6.3.3 Strictly anaerobic fermentation with growth factors ergosterol and Tween 80

Ergosterol and Tween 80 are important anaerobic growth factors [117-119], and are routinely used under strictly anaerobic conditions where culture is sparged with pure nitrogen. To testify the ability of H153E10-XYLA31 to ferment under strictly anaerobic

condition, and also to verify the effect of anaerobic growth factors ergosterol and Tween 80, the fermentation profile of H153E10-XYLA31 with the two anaerobic growth factors and sparged with N_2 was compared with fermentation results without the factors and N_2, as shown in Figure 6-5 and Table 6-5.

With the anaerobic growth factors, the specific growth rate and biomass yield were both about 20% higher. In addition, the overall xylose consumption rate showed a 30% improvement. As a result, the duration of the fermentation was reduced by 20%. Sparging with nitrogen didn't significantly affect the growth and fermentation, but caused lower ethanol yield due to the ethanol evaporation and inefficient condensation. Like previous fermentation results of H153E10-XYLA31 on 4% xylose, the xylose was depleted when the cell was still in growth phase, and hence limited the growth.

To address the above issues, 80 g/L xylose was used to initiate the fermentation. Also, the anaerobic conditions in our previous fermentation setup were rather strict since cultivation were purged with ultra-pure N_2 at the beginning of each fermentation, and the diffusion of O_2 into the system was minimal due to the slightly positive pressure in the fermenter caused by CO_2 production. Thus, the fermentation profile of H153E10-XYLA31 with the two anaerobic growth factors was compared with fermentation results without the factors (from Figure 6-4 and Table 6-4), and combined into Figure 6-6 and Table 6-6.

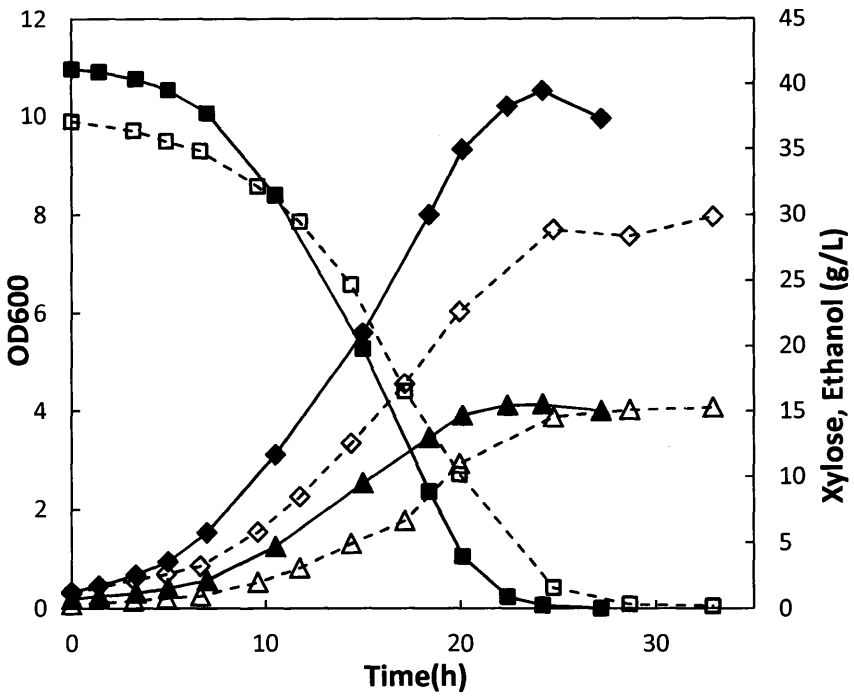


Figure 6-5. Effect of strictly anaerobic condition with ergosterol and Tween 80

With N_2 sparging, ergosterol (10mg/L) and Tween 80 (0.42g/L): (◆OD_{600}, ■Xylose, ▲Ethanol)

Without N_2 sparging, ergosterol or Tween 80: (◇OD_{600}, □Xylose, △Ethanol)

Table 6-5. Characteristics of H153E10-XYLA31 fermentation under strictly anaerobic condition with ergosterol and Tween 80

		No N_2 sparging or growth factors		N_2 sparging, 10mg/L ergosterol and 0.42g/L Tween80	
		Growth phase	Overall	Growth phase	Overall
μ_{max}	h^{-1}	0.177±0.009	/	0.215±0.004	/
Δt	h	19.85	29.70	20.10	24.20
Δ[Xylose]	g/L	34.02	36.19	36.98	40.91
Δ[Ethanol]	g/L	13.75	14.69	13.76	14.80
Δ[DCW]	g/L	1.869	1.968	2.363	2.717
Δ[Glycerol]	g/L	2.114	2.134	2.512	2.670
Final OD_{600}		\	7.96	\	10.52
$OD_{average}$		3.780	4.675	3.832	4.945
$DCW_{average}$	g/L	1.008	1.336	1.022	1.413
q_{xylose}	$g \cdot L^{-1} h^{-1}$	1.714	1.219	1.840	1.690
	$g \cdot g^{-1} h^{-1}$	1.700	0.912	1.801	1.196
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	0.693	0.495	0.684	0.611
	$g \cdot g^{-1} h^{-1}$	0.687	0.370	0.670	0.433
Ethanol yield	g/g	0.404	0.406	0.372	0.362
Biomass yield	g/g	0.055	0.054	0.064	0.066
Glycerol yield	g/g	0.062	0.059	0.068	0.065

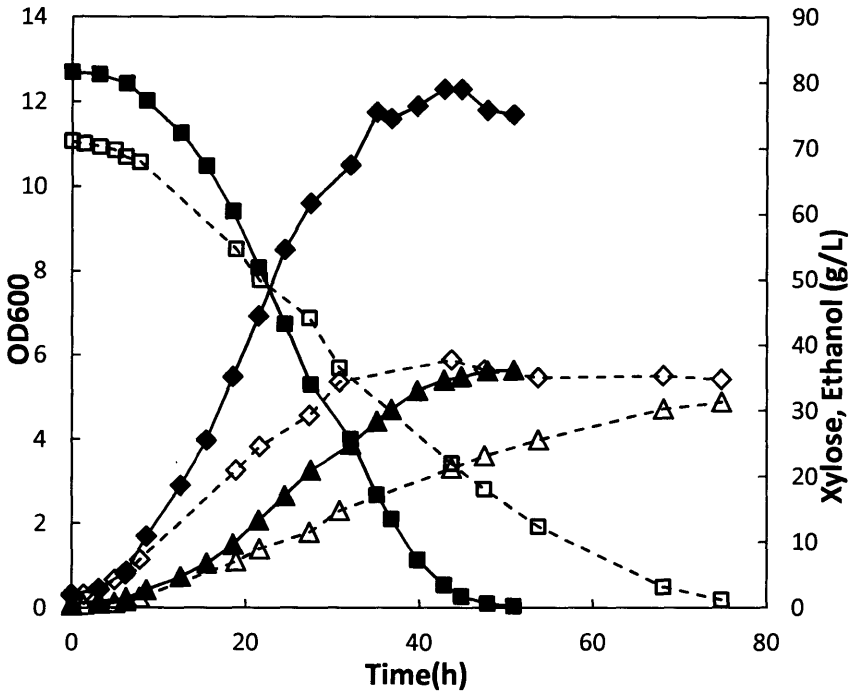


Figure 6-6. Effect of ergosterol and Tween 80 on H153E10-XYLA31 fermentation

With ergosterol (10mg/L) and Tween 80 (0.42g/L): (◆OD_{600}, ■Xylose, ▲Ethanol)

Without ergosterol or Tween 80: (◇OD_{600}, □Xylose, △Ethanol)

Table 6-6. Characteristics of H153E10-XYLA31 fermentation with ergosterol and Tween 80

		No ergosterol and Tween80		10mg/L ergosterol and 0.42g/L Tween80	
		Growth phase	Overall	Growth phase	Overall
μ_{max}	h^{-1}	0.202±0.009	/	0.199±0.010	/
Δt	h	24.15	74.80	24.40	47.85
Δ[Xylose]	g/L	26.08	69.94	47.20	81.06
Δ[Ethanol]	g/L	10.83	30.99	20.00	35.47
Δ[DCW]	g/L	1.103	1.369	2.441	3.061
Δ[Glycerol]	g/L	1.530	2.799	2.546	3.520
Final OD_{600}		\	5.420	\	11.7
$OD_{average}$		2.547	4.332	4.403	7.594
$DCW_{average}$	g/L	0.679	1.155	1.174	2.025
q_{xylose}	$g \cdot L^{-1} h^{-1}$	1.080	0.935	1.935	1.694
	$g \cdot g^{-1} h^{-1}$	1.590	0.809	1.647	0.837
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	0.448	0.414	0.820	0.741
	$g \cdot g^{-1} h^{-1}$	0.660	0.359	0.698	0.366
Ethanol yield	g/g	0.415	0.443	0.424	0.438
Biomass yield	g/g	0.042	0.020	0.052	0.038
Glycerol yield	g/g	0.059	0.040	0.054	0.043

With the supplement of ergosterol and Tween 80, H153E10-XYLA31 was able to finish 81 g/L xylose in about 48 hours, significantly faster than the same strain without the growth factors (70 g/L xylose in 75 hours). The initial growth rates, specific xylose consumption rates and ethanol productivities are almost identical for strain under the two conditions, as well as the yields of ethanol and glycerol. The only difference is the nearly twice higher biomass production with growth factors added, which led to about 80% higher volumetric xylose consumption rate and ethanol productivity. Xylose was no longer a limit of growth since the culture entered the stationary phase when a sufficient amount of xylose remained.

6.3.4 Expression of other heterologous XIs

Expression of quite a few heterologous XIs from various organisms had been attempted in *S. cerevisiae*, with very limited success[45]. The successful expression of *Piromyces XYLA* in *S. cerevisiae* [47] is a milestone after 25 years of elaboration. XI from anaerobic fungus *Orpinomyces* [22], and bacterium *Clostridium phytofermentans* [24] have also been successfully expressed in *S. cerevisiae.* However, all the XIs could only enable a very low initial growth rate and further metabolic and evolutionary engineering were required to increase the xylose utilization. According to the current understanding of the XI expression, it is promising to revisit some of the unsuccessfully expressed XIs, adding another powerful tool of codon/expression optimization for *S. cerevisiae.*

The XIs from *E.coli, Bacillus subtilis, Bacteroides thetaiotaomicron, Thermus thermophiles, Streptomyces rubiginosus, Arabidopsis thaliana* and *Xanthomonas*

campestris (see Appendix for sequences) were picked for expression in *S. cerevisiae*, to represent each cluster shown in Figure 6-7.

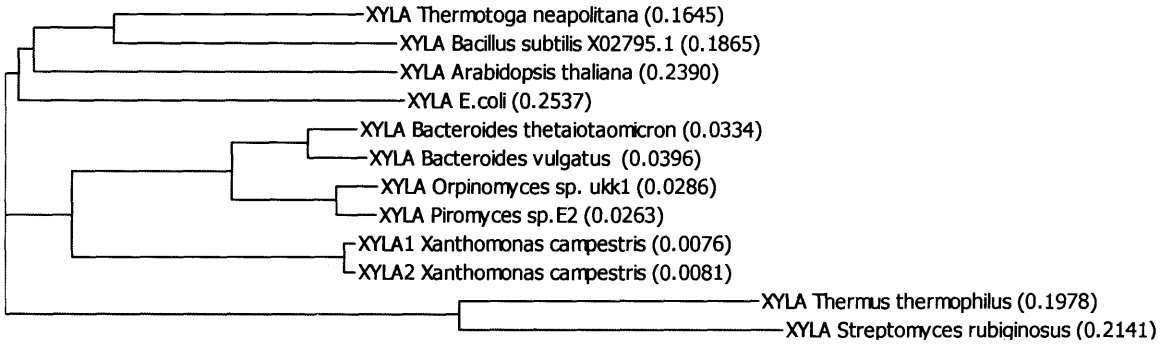


Figure 6-7. Clustering of XI

All the genes are codon/expression optimized and synthesized before being cloned to multiple copy plasmid pRS426HXT7$_{Tr}$-p (Table 6-1). The constructed plasmids were then transformed to H142 (PPP and *XYL3* overexpressed) and evolutionarily engineered through aerobic sequential batch cultivation in a rich medium, similar to the process described in Section 5.3.1.

After about 4 weeks of evolution, only the *Bacteroides thetaiotaomicron XYLA* showed adequate XI activity to support growth on xylose, other than previously proved *Piromyces XYLA*. The aerobic growth profile of H142E-BthXYLA compared with H142E-PXYLA and H131E1-XYLA31 is shown in Figure 6-8.

The newly evolved H142E-BthXYLA showed similar aerobic growth as H142E-PXYLA (newly evolved strain as a control), and slightly better than the previously engineered H131E1-XYLA31. Not surprisingly, the *B. thetaiotaomicron XYLA* shares about 80% amino acid sequence homology with *Piromyces XYLA*. A qPCR revealed

31.2±2.7 copies of *BthXYLA* in H142E-BthXYLA, implying that *BthXYLA* also needs to be overexpressed to support growth of the host on xylose, similar to the *Piromyces XYLA*.

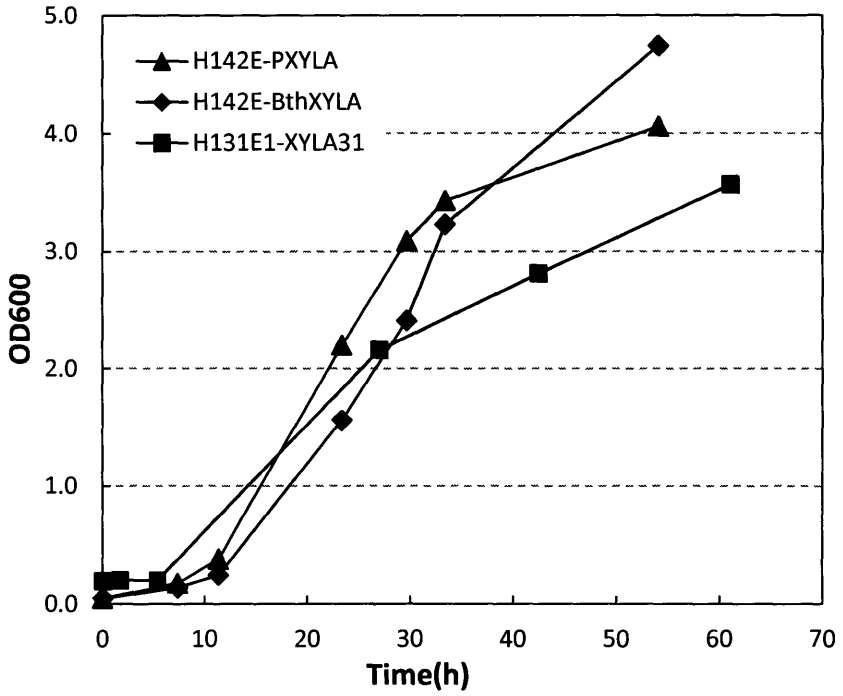


Figure 6-8. Aerobic growth of *S. cerevisiae* containing *Bacteroides thetaiotaomicron XYLA*

6.3.5 NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (*gapN*)

A previous study had shown that anaerobic growth factor ergosterol can greatly boost growth on xylose medium, as well as more complex media with yeast extracts. These observations suggest insufficient biosynthesis of certain factors under anaerobic condition. One hypothesis is that the high flux of xylose catabolism saturates the non-oxidative branch of the pentose phosphate pathway, and therefore represses the oxidative PPP where cytosol NADPH is produced. A lack of NADPH can cause deficiency of amino acid and steroid (like ergosterol) synthesis and consequently slower growth, which could be alleviated by supplementing with ergosterol or yeast extract.

To test this hypothesis, a NADP-dependent non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase (*gapN*) from *Streptococcus mutants* was expressed in H131E5-XYLA31. The *gapN* pathway is shown in Figure 6-9.

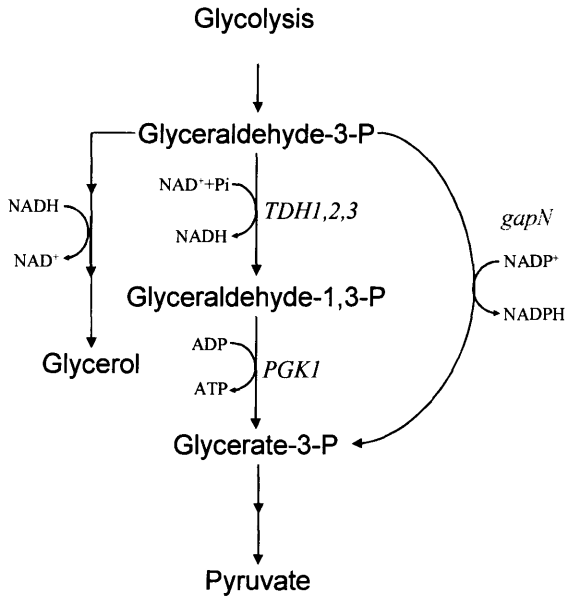


Figure 6-9. NADP-dependent non-phosphorylating G3P dehydrogenase pathway

It has been reported that by diverting some G3P from wild-type GAPD pathway, the cell could make more NADPH and less NADH and ATP, consequently result in 50% less glycerol and 3% more ethanol [120].

gapN from *Streptococcus mutants* was cloned into YIp pUCAR1, resulting in pUCAR1-gapN (Table 6-1). H131E5-XYLA31 was then transformed with *Hpa* I digested pUCAR1-gapN to make strain H141E5-XYLA31 (Table 6-2). H131E5-XYLA31 transformed with empty pUCAR1 was used as control. The fermentation profiles of the two strains are shown in Figure 6-10.

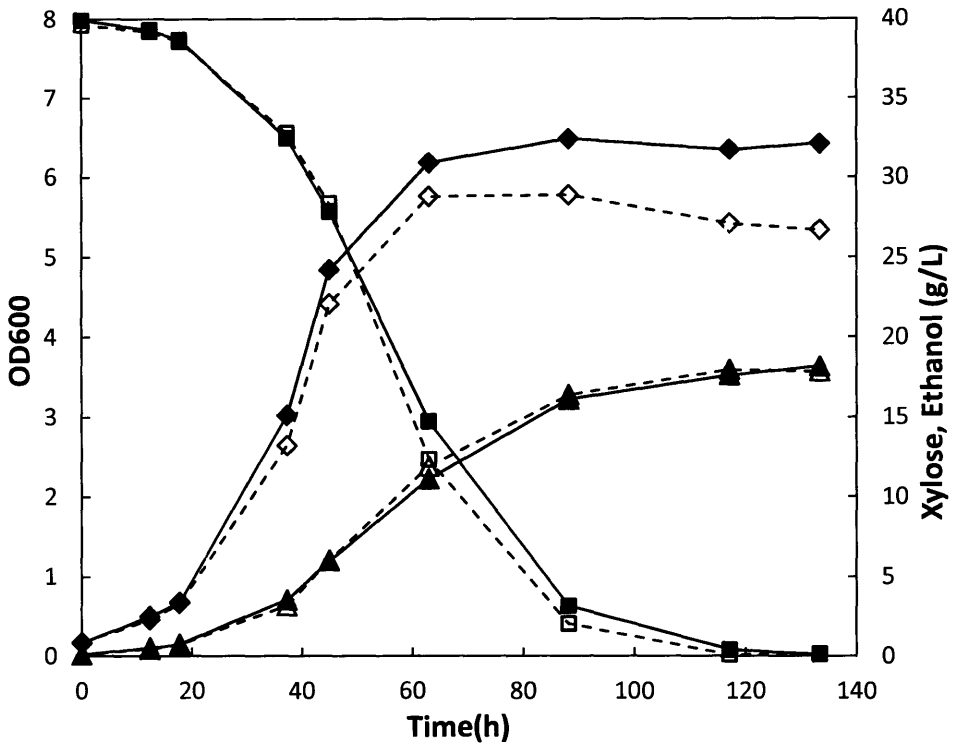


Figure 6-10. Effect of *gapN* on xylose fermentation

H141E5-XYLA31 (gapN): (◆OD_{600}, ■Xylose, ▲Ethanol)

H131E5-XYLA31-ARG4 (control): (◇OD_{600}, □Xylose, △Ethanol)

Table 6-7. Characteristics of H141E5-XYLA31 fermentation

		H131E5-XYLA31-ARG4 Control	H141E5-XYLA31 *gapN*
μ_{max}	h^{-1}	0.073±0.002	0.077±0.003
Δt	h	117.15	117.15
Δ[Xylose]	g/L	39.54	39.54
Δ[Ethanol]	g/L	17.85	17.51
Δ[DCW]	g/L	1.399	1.646
Δ[Glycerol]	g/L	1.949	1.485
Final OD_{600}		5.420	6.340
$OD_{average}$		3.974	4.423
$DCW_{average}$	g/L	1.060	1.179
q_{xylose}	$g \cdot L^{-1} h^{-1}$	0.338	0.337
	$g \cdot g^{-1} h^{-1}$	0.319	0.286
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	0.152	0.149
	$g \cdot g^{-1} h^{-1}$	0.144	0.127
Ethanol yield	g/g	0.451	0.443
Biomass yield	g/g	0.035	0.042
Glycerol yield	g/g	0.049	0.038

The expression of *gapN* provided marginal improvement on xylose fermentation, reducing glycerol production by 25% and increasing final biomass by 20%, but didn't change xylose consumption rate or ethanol productivity significantly. pUCAR1-gapN was also transformed to the better evolved H141E10-XYLA31, but again no significant improvement of xylose fermentation could be observed, although glycerol production was reduced by 20% (data not shown).

6.3.6 Expression of heterologous xylose transporters

As discussed in Chapter 1, inefficient xylose transport could be a limiting step in xylose assimilation by *S. cerevisiae*. Xylose utilization has been greatly improved after successful metabolic engineering, which makes the xylose transport more likely the bottleneck of the system.

Twelve putative transporters from 5 natural xylose-utilizing strains (*Candida albicans, Candida glabrata, Debaryomyces hansenii, Pichia stipitis* and *Trichoderma reesei*) were selected according to sequence homology. They were separately cloned into shuttle vector pRS424 (Table 6-1) and the plasmids were transformed into yeast. Before the H131 strain was successfully constructed, the putative xylose transporters were tested in a legacy strain YSX3 (Table 6-2).

The 12 strains together with the control strain underwent a sequential batch cultivation on selective SD media for 3 rounds, and 1% xylose was used instead of regular 2% to further increase the selection pressure. After the selection, the mixed culture was diluted appropriately and plated on SDX (1%) for screening. The larger colonies were picked to segregate the plasmids for transporters sequencing. The result was confirmed by retransforming plasmids to YSX3, and two transformants containing

DhSUT1 (from *Debaryomyces hansenii*) and *CgSUT1* (from *Candida glabrata*) displayed an advantage in aerobic growth on xylose, as shown in Figure 6-11.

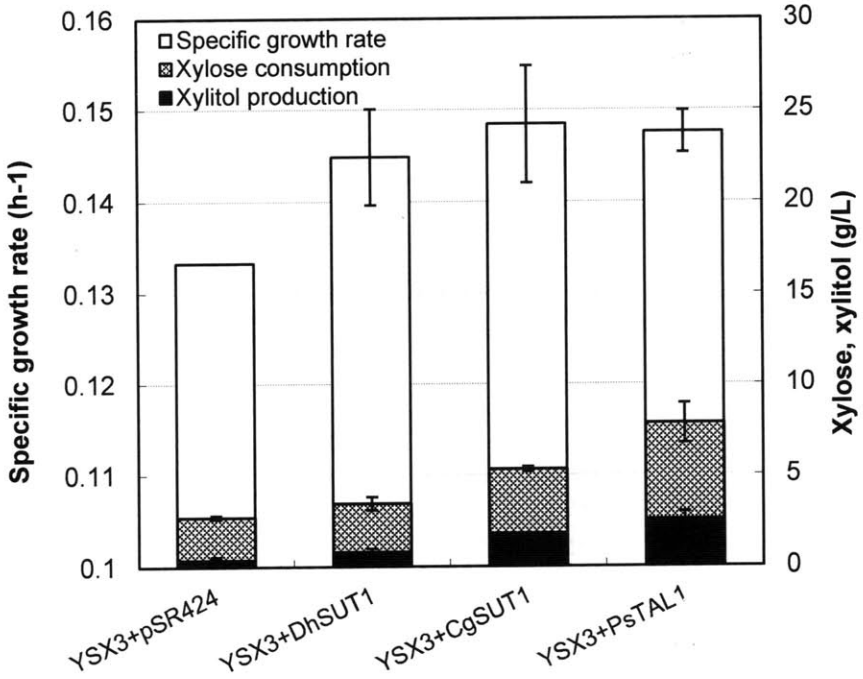


Figure 6-11. Performance of strain with selected transporters

The YSX3 transformed with *DhSUT1* and *CgSUT1* showed 2-4 times higher xylitol production and 30-90% higher xylose consumption than the control strain. The contribution of the transporters on xylose consumption was not as significant as that of *PsTAL1* (a PPP gene). The effect of the transporters is to be further validated in the newly engineered H153E10-XYLA31, in which the limit of xylose transport could be more prominent.

6.3.7 Chemostat cultivation of H153E10-XYLA31

To explore the maximum fermentation capacity of the strain as well as the maximum specific growth rate, strain H153E10-XYLA31 underwent chemostat in minimal media with sufficient xylose. The bioreactor setup was as described previously,

with the addition of anaerobic growth factors ergosterol and Tween 80. The dilution rate was increased stepwise by 0.02 h^{-1}, once the chemostat reached a steady state. The results at dilution rate of 0.1-0.2 h^{-1} are shown in Table 6-8.

Table 6-8. Characteristics of H141E5-XYLA31 in chemostat

D	h^{-1}	0.098	0.118	0.137	0.157	0.176	0.196
$[Xylose]_{Feed}$	g/L	52.692	45.392	44.472	42.951	41.089	42.723
$[Xylose]_{Effluent}$	g/L	23.090	25.122	29.159	34.362	33.856	38.629
Δ[Xylose]	g/L	29.602	20.269	15.313	8.589	7.233	4.094
$[Ethanol]_{Feed}$	g/L	0.708	0.708	0.708	0.758	0.854	0.628
$[Ethanol]_{Effluent}$	g/L	14.251	9.243	6.541	3.765	3.246	1.969
Δ[Ethanol]	g/L	13.543	8.535	5.833	3.007	2.392	1.341
Δ[Glycerol]	g/L	1.350	0.951	0.713	0.422	0.371	0.139
OD_{600}		4.120	3.193	2.655	1.797	1.490	0.978
DCW	g/L	1.099	0.851	0.708	0.479	0.397	0.261
q_{xylose}	$g \cdot L^{-1} h^{-1}$	2.915	2.391	2.104	1.347	1.275	0.801
	$g \cdot g^{-1} h^{-1}$	2.654	2.808	2.972	2.812	3.210	3.075
$q_{ethanol}$	$g \cdot L^{-1} h^{-1}$	1.334	1.007	0.801	0.472	0.422	0.262
	$g \cdot g^{-1} h^{-1}$	1.214	1.182	1.132	0.984	1.062	1.007
Ethanol yield	g/g	0.458	0.421	0.381	0.350	0.331	0.328
Biomass yield	g/g	0.037	0.042	0.046	0.056	0.055	0.064
Glycerol yield	g/g	0.046	0.047	0.047	0.049	0.051	0.034

At dilution rate 0.1 h^{-1}, a significant amount (56%) of xylose in feed was consumed to produce ethanol at a very good yield (0.456 g/g). The specific xylose consumption rate and ethanol productivity reached 2.654 $g \cdot g^{-1} h^{-1}$ and 1.214 $g \cdot g^{-1} h^{-1}$, respectively, which is about 60% better than the performance in batch cultivation and is the highest ever observed in the research. With the increase of the dilution rate, the specific xylose consumption rate and ethanol productivity were even higher. Nevertheless, the consumed xylose and ethanol production dropped dramatically due to the rapid decline of biomass. In addition, the ethanol yield also fell monotonically with increasing

biomass yield. The results diminished the practical significance of the high specific fermentation rate in the chemostat cultivation.

6.3.8 Discussion

As summarized in Table 6-9, the final metabolic and evolutionary engineered strain H153E10-XYLA31 showed a maximum anaerobic growth rate of 0.199 h^{-1}, 65% better than that of H131E8-XYLA31 before optimization or the best reported strain RWB 218. Xylose consumption rate of H153E10-XYLA31 was 1.647 $g \cdot g^{-1}h^{-1}$ in batch fermentation, 80% higher than before optimization and 37% higher than RWB 218, with same level of ethanol yield (data were taken from the growth phase to match the condition in the literature). In chemostat cultivation of H153E10-XYLA31, the peak xylose consumption rate reached 2.654 $g \cdot g^{-1}h^{-1}$ at a dilution rate of 0.1 h^{-1}, 60% higher than in the batch cultivation of the same strain.

The evolved strain performs well at elevated sugar concentrations (80 g/L); however the biomass yield was still relatively low (0.038 g/g) when the carbon source was not limiting. A similar final OD_{600} of around 10 was also achieved under aerobic and anaerobic condition, implying that the anaerobic growth factors ergosterol and Tween 80 are not the limiting factors any more. Additional optimization of the medium (such as phosphate and nitrogen source) and metabolic engineering of the strain toward better tolerance to stresses may be necessary to further improve the growth and fermentation.

The fermentation performance of strain H153E10-XYLA31 is superior to any other reported strain to the best of our knowledge. Further metabolic engineering toward mixed sugar fermentation and successful transfer of fermentation performance to industrial strains and conditions should be the key issues to stress in future research.

Table 6-9. Characteristics of recombinant *S. cerevisiae* strains

Strain	Strain description	Fermentation condition	Yields g/g		$q_{Ethanol}$ $g \cdot g^{-1} h^{-1}$	q_{Xylose} $g \cdot g^{-1} h^{-1}$	μ_{max} h^{-1}
			Ethanol	Xylitol			
H131-XYLA31	*XYLA, PsXYL3, PsTAL1, TKL1, RPE1, RKI1*	Aerobic batch, SDX, 2% xylose	N/A*	N/A	N/A	N/A	0.031±0.022
H131E1-XYLA31	Selection of H131-XYLA31, aerobic sequential batch	Micro-aerobic batch, SDX, 2% xylose	0.273	<0.01	0.041	0.148	0.197±0.006
H131E3-XYLA31	Selection of H131-XYLA31, micro-aerobic sequential batch	Anaerobic batch, 2×YNB, 4% xylose	0.417	<0.01	0.122	0.293	0.061±0.002
H131E5-XYLA31	Selection of H131-XYLA31, anaerobic sequential batch	Anaerobic batch, 2×YNB, 4% xylose	0.410	<0.01	0.269	0.655	0.073±0.002
H131E8-XYLA31	Selection of H131-XYLA31, xylose-limited anaerobic chemostat	Anaerobic batch, 2×YNB, 4% xylose	0.423	<0.01	0.387	0.916	0.120±0.004
		Anaerobic chemostat, 1×YNB, 1.5% xylose	0.438	<0.01	0.687	1.568	0.148
H153E10-XYLA31	H131E8-XYLA31, with complement of *ARG4* and *LEU2*	Anaerobic batch, 2×YNB, 8% xylose, ergosterol and Tween 80	0.424	<0.01	0.698	1.647	0.199±0.010
		Anaerobic chemostat, 2×YNB, 4% xylose, ergosterol and Tween 80	0.458	<0.01	1.214	2.654	0.098
RW 202-AFX[19]	*Piromyces XYLA* evolved isolate	Anaerobic batch, synthetic medium	0.42	0.021	0.14	0.34	0.03
RWB 217[20]	*PXYLA, XKS1, TAL1, TKL1, RPE1, RKI1, Δgre3*	Anaerobic batch, synthetic medium	0.43	0.003	0.46	1.06	0.09
RWB 218[3, 21]	Selection of RWB 217	Anaerobic batch, synthetic medium	0.41	0.001	0.49	1.2	0.12

*N/A = not tested

6.4 Conclusions

The restoration of the auxotrophic markers *arg4* and *leu2* (primarily *leu2*) in H131E10-XYLA31 greatly boosted aerobic growth on minimal medium. The anaerobic growth can be further improved by supplementing the anaerobic growth factors (ergosterol and Tween 80) or complex medium gradient yeast extract. Under the optimized conditions, the engineered strain H153E10-XYLA31 showed a anaerobic μ_{max} of 0.199 h^{-1} and specific xylose consumption rate of 1.647 $g \cdot g^{-1} h^{-1}$ in batch fermentation, 65% and 37% higher than those of the best reported xylose-fermenting strain RWB 218, respectively. In chemostat cultivation, the specific performance of H153E10-XYLA31 was 60% better than that of the same strain in the growth phase of batch cultivation, although a significant amount of xylose was left unconsumed in the effluent.

The strictly anaerobic condition realized by sparging ultra-pure nitrogen didn't affect the fermentation performance, although the ethanol yield was leveled down due to evaporation. Expression of NADP-dependent G3P dehydrogenase (*gapN*) was able to reduce glycerol production by 25%, but led to only marginal improvement in anaerobic growth. Xylose isomerase from bacterium *Bacteroides thetaiotaomicron* was successfully expressed in *S. cerevisiae*, resulting in a xylose-utilizing strain equivalent to similarly constructed H131E1-XYLA31. Heterologous expression of putative xylose transporters (SUTs) showed a promising result in a host (YSX3) far inferior to the newly engineered strain H153E10-XYLA31. Therefore, more significant results can be expected through further study of the SUTs in H153E10-XYLA31, where sugar transport is more likely the bottleneck of xylose assimilation.

CHAPTER 7. Conclusions and Recommendations

As discussed throughout this work, a central goal of this thesis has been to enable and improve the utilization of xylose by *S. cerevisiae*, for production of bioethanol and potentially other chemicals. To achieve superior xylose-fermenting strains, both rational and combinatorial metabolic engineering approaches were applied for the construction, adaption and characterization of the strains.

7.1 Conclusions

Xylose-utilizing strains were successfully constructed through the heterologous expression of either the XI-XK (*XYLA-XYL3*) or the XR-XDH-XK (*XYL1-XYL2-XYL3*) xylose metabolic pathways as well as the non-oxidative pentose phosphate pathway, resulting in strains H131-XYLA31 and H131-XYL123, respectively. Both strains showed slow but significant aerobic growth (μ_{max}= 0.031±0.022 h^{-1} for H131-XYLA31 and 0.081±0.052 $h^{-1}$for H131-XYL123) after pathway construction.

The engineered strain carrying the redox XR-XDH pathway showed substantial accumulation of the by-product, xylitol, due to the intrinsic cofactor imbalance of the pathway. Therefore, evolutionary engineering approaches were put into practice mainly on the XI-based strain H131-XYLA31 for improved xylose utilization. After multiple stages of evolution including aerobic to oxygen-limited sequential batch cultivation, followed by xylose-limited anaerobic chemostat cultivation, both the anaerobic growth rate and the xylose consumption rate of the evolved strain were significantly improved.

The finally isolated strain, H131E8-XYLA31, rapidly grew and fermented xylose under anaerobic conditions, with a specific growth rate of 0.12 h^{-1} and a specific xylose consumption rate exceeding 0.9 $g \cdot g^{-1} h^{-1}$. The evolutionary engineering approach was again proven to be a powerful tool for obtaining improved properties in the target strain, such as efficient substrate utilization in this study.

Further modification of restoring the auxotrophic markers *arg4* and *leu2* (primarily *leu2*) in H131E10-XYLA31 greatly boosted its aerobic growth on a minimal medium. In addition, the anaerobic growth can be dramatically improved by supplementing the anaerobic growth factors, ergosterol and Tween 80. After optimization of the metabolically and evolutionarily engineered strain, H153E10-XYLA31, an anaerobic specific growth rate of 0.199 h^{-1} and a specific xylose consumption rate of 1.647 $g \cdot g^{-1} h^{-1}$ were achieved in batch fermentation, 65% and 37% higher than those of the best reported xylose-fermenting strain RWB 218 [3, 21], respectively. The ethanol yield was as high as 0.438 g/g and the xylitol accumulation was negligible. In addition, the specific performance was 60% better in chemostat cultivation than in the growth phase of batch cultivation, although a significant amount of xylose left unconsumed in the effluent.

A successful inverse metabolic engineering approach was demonstrated to identify gene targets responsible for the phenotype of rapid xylose assimilation. A genomic library of the evolutionarily engineered H131E5-XYLA31 was constructed and amplified in *E. coli*, then transformed to the *S. cerevisiae* host for screening. A high throughput screening system is the key factor of such inverse metabolic engineering. The microfluidic sorting system aiming for quick xylose consumption is much more sensitive than the screening based on growth rate, despite lower coverage of the library. The main

genotype identified in the evolved strain is the duplication of *XYLA*. The genotype was verified by the structural analysis of the plasmid selected after the screening, and qPCR of the genomic DNA from the evolved strain, H131E5-XYLA31.

A high expression level of XI was confirmed by molecular biology techniques including quantitative PCR, blotting and hybridization and enzyme activity assay. The tandem gene duplication (TGD) of *XYLA* was proved in the series of evolved H131E-XYLA31 strains. And the TGD structure was translocated onto one or more chromosomes during the early phase of the evolution. However, the TGD is not the only way to achieve high XI expression. The evolution process was successfully reproduced in a different background H142 where *XYLA* expansion could originate from different initial *XYLA* construction, whether through chromosomal integration or on different types of plasmids. The site-specific *XYLA* integration and duplication could also be achieved, although the structure might be subject to other genomic rearrangement.

The XI from bacterium *Bacteroides thetaiotaomicron* was successfully expressed in *S. cerevisiae*, resulting in a xylose-utilizing strain equivalent to the similarly constructed and evolutionary engineered H131E1-XYLA31. A XI with higher specific activity in *S. cerevisiae* is still desirable, yet none of other XIs expressed in *S. cerevisiae* showed significantly better specifications.

The XK (*XYL3*) form *P. stipitis* could be efficiently expressed in *S. cerevisiae*, and low copy number of *PsXYL3* promoted by strong promoter should be sufficient to support a high flux of xylose assimilation.

Mutations on *GRE3* that potentially abolish XR activity were discovered in the evolutionarily engineered strains, supporting the hypothesis that the elimination of XR

activity was favorable for xylose utilization by reducing the production of xylitol and hence alleviating the inhibition of XI activity caused by xylitol.

The strictly anaerobic conditions realized through sparging ultra-pure nitrogen did not affect the fermentation performance, although the ethanol yield was leveled down due to ethanol evaporation caused by gas flow.

The expression of an NADP-dependent G3P dehydrogenase (*gapN*) was able to reduce glycerol production by 25%, but led to only marginal improvement in anaerobic growth.

Heterologous expression of putative xylose transporters (SUTs) showed promising results in a host (YSX3) with much lower xylose assimilation rate than the newly engineered H153E10-XYLA31. Therefore, more significant results can be expected through further study of the SUTs in H153E10-XYLA31, where sugar transport is more likely the bottleneck of xylose metabolism.

The fermentation performance of the strain H153E10-XYLA31 is superior to any other reported strain to the best of our knowledge. Further metabolic engineering toward successful transfer of fermentation performance to industrial strains and conditions should be the key issue to stress in future research.

7.2 Recommendations and future work

After intensive metabolic engineering, the xylose fermentation performance of the strain H153E10-XYLA31 has been greatly improved and is no longer a true bottleneck in the fermentation of hemicellulose hydrolysates. However, this does not imply that further improvement is either impossible or undesirable.

As we proposed and preliminarily tested, the xylose transport may remain a rate-limiting step of xylose assimilation, since the downstream pathway has been adequately engineered. Particularly, at low extracellular xylose concentration or when fermenting on mixed sugar, a high-specific xylose transport system will be necessary for efficient sugar utilization. Expression of heterologous (putative) xylose transporters from native xylose-utilizing strains (preferable fungi) is still the most feasible strategy, despite the limited success after decades of effort. Upon the discovery of effective xylose transporters, more engineering such as mutagenesis and gene shuffling could be used to further improve the capacity and specificity of the xylose transporters. In addition, extensive evolutionary engineering may be another potential approach to adapt and enhance the endogenous sugar transporters, when carried out using a low xylose concentration as the selective pressure.

Mixed sugar fermentation is another challenge for the rapid and complete utilization of hemicellulose hydrolysates, mainly due to the effect of Carbon Catabolite Repression (CCR). Expression of a xylose-specific transport may alleviate the CCR and therefore promote the co-utilization of hexoses and pentoses. Also, evolutionary engineering may be required to improve the strains in terms of mixed sugars fermentation.

The XIs from *Piromyces* and some other fungi or bacteria have been proved to support efficient xylose assimilation. However, a very high expression level of the enzyme is required, implying a low specific activity of the XIs expressed in *S. cerevisiae*. Therefore, efforts either to search for XIs with better properties or to engineer the current XIs could be favorable to higher xylose-assimilation efficiency and lower metabolic burden of abundant XI synthesis.

The expression of an NADP-dependent G3P dehydrogenase (*gapN*) showed reduced glycerol production and marginal improvement in anaerobic growth, when integrated into the chromosome and expressed in parallel with the endogenous G3P dehydrogenase (*TDH1,2,3*). It can be expected that a fine-tuned expression of *gapN* together with *TDH* may further change the NADPH/NADH ratio, resulting in better growth and lower concentration of byproduct glycerol.

The inverse metabolic engineering approach serves as a useful tool to identify the genotypes of an evolutionarily engineered strain; however, two obstacles limited its application in current research. The first was the difficulty in construction of a knockout yeast library. The genomic library could only represent the dominant mutations, while for identification of recessive phenotypes a knockout library is required. Knockout libraries can be efficiently generated using transposon in bacteria, but hardly in yeast. It is possible to prepare and fragment the genomic DNA of target strain, apply the transposon insertion *in vitro* and retransform the DNA back to the target strain to generate a knockout library. However, the efficiency of transposon insertion and transformation must be remarkably improved for good quality of the library. The second obstacle was the throughput of the microfluidic screening method. The microfluidic device which coalesces, detects, and sorts droplets is currently the throughput bottleneck since it can only sort up to 10^4 droplets in one hour while the droplet formation rate of the cell encapsulation device is 10^6 droplets in the same amount of time. By optimizing the sorting device, libraries with sizes of 10^5 or higher could be screened, which represents a better coverage of the yeast genome for an overexpression or knockout library.

References

[1] M. L. Khandekar, *et al.*, "The global warming debate: A review of the state of science," *Pure and Applied Geophysics,* vol. 162, pp. 1557-1586, Aug 2005.

[2] M. Wang, *et al.*, "Effects of Fuel Ethanol Use on Fuel-Cycle Energy and Greenhouse Gas Emissions," Center for Transportation Research, Energy Systems Division, Argonne National Laboratory1999.

[3] A. J. a. van Maris, *et al.*, "Alcoholic fermentation of carbon sources in biomass hydrolysates by Saccharomyces cerevisiae: current status," *Antonie van Leeuwenhoek,* vol. 90, pp. 391-418, 2006.

[4] V. Smil, "Crop residues: agriculture's largest harvest," *BioScience,* vol. 49, pp. 299-308, 1999.

[5] C. E. Wyman, "BIOMASS ETHANOL: Technical Progress, Opportunities, and Commercial Challenges," *Annual Review of Energy and the Environment,* vol. 24, pp. 189-226, 1999.

[6] H. G. Lawford and J. D. Rousseau, "Production of ethanol from pulp mill hardwood and softwood spent sulfite liquors by genetically engineered E. coli," *Appl Biochem Biotechnol,* vol. 39-40, pp. 667-85, Spring 1993.

[7] T. W. Jeffries and Y. S. Jin, "Metabolic engineering for improved fermentation of pentoses by yeasts," *Applied microbiology and biotechnology,* vol. 63, pp. 495-509, 2004.

[8] J. C. d. Preez, *et al.*, "D-xylose fermentation by Candida shehatae and Pichia stipitis at low dissolved oxygen levels in fed-batch cultures," *Biotechnology Letters,* vol. 11, pp. 131-136, 1989.

[9] T. W. Jeffries, "Conversion of xylose to ethanol under aerobic conditions by Candida tropicalis," *Biotechnology letters,* vol. 3, pp. 213-218, 1981.

[10] H. B. Klinke, *et al.*, "Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass," *Appl Microbiol Biotechnol,* vol. 66, pp. 10-26, Nov 2004.

[11] W. Visser, *et al.*, "Oxygen requirements of yeasts," *Appl Environ Microbiol,* vol. 56, pp. 3785-92, Dec 1990.

[12] N. W. Ho, *et al.*, "Genetically engineered Saccharomyces yeast capable of effective cofermentation of glucose and xylose," *Applied and environmental microbiology,* vol. 64, pp. 1852-9, 1998.

[13] N. W. Y. Ho, *et al.*, "Genetically Engineered Saccharomyces Yeasts for Conversion of Cellulosic Biomass to Environmentally Friendly Transportation Fuel Ethanol," in *ACS Symposium Series* ed: American Chemical Society, 2000, pp. 143-159.

[14] M. Jeppsson, *et al.*, "The level of glucose-6-phosphate dehydrogenase activity strongly influences xylose fermentation and inhibitor sensitivity in recombinant Saccharomyces cerevisiae strains," *Yeast,* vol. 20, pp. 1263-72, Nov 2003.

[15] M. Jeppsson, *et al.*, "The expression of a Pichia stipitis xylose reductase mutant with higher KM for NADPH increases ethanol production from xylose in recombinant Saccharomyces cerevisiae," *Biotechnology and bioengineering,* vol. 93, pp. 665-73, 2006.

[16] K. Karhumaa, *et al.*, "Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant Saccharomyces cerevisiae," *Microbial cell factories,* vol. 6, pp. 5-5, 2007.

[17] Y.-s. Jin, *et al.*, "Optimal Growth and Ethanol Production from Xylose by Recombinant Saccharomyces cerevisiae Require Moderate D -Xylulokinase Activity," *APPLIED AND ENVIRONMENTAL MICROBIOLOGY,* vol. 69, pp. 495-503, 2003.

[18] Y. S. Jin, *et al.*, "Improvement of xylose uptake and ethanol production in recombinant Saccharomyces cerevisiae through an inverse metabolic engineering approach," *Applied and environmental microbiology,* vol. 71, pp. 8249-8249, 2005.

[19] M. Kuyper, *et al.*, "Minimal metabolic engineering of Saccharomyces cerevisiae for efficient anaerobic xylose fermentation: a proof of principle," *FEMS yeast research,* vol. 4, pp. 655-64, 2004.

[20] M. Kuyper, *et al.*, "Metabolic engineering of a xylose-isomerase-expressing Saccharomyces cerevisiae strain for rapid anaerobic xylose fermentation," *FEMS yeast research,* vol. 5, pp. 399-409, 2005.

[21] M. Kuyper, *et al.*, "Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting Saccharomyces cerevisiae strain," *FEMS yeast research,* vol. 5, pp. 925-34, 2005.

[22] A. Madhavan, *et al.*, "Xylose isomerase from polycentric fungus Orpinomyces: gene sequencing, cloning, and expression in Saccharomyces cerevisiae for bioconversion of xylose to ethanol," *Applied microbiology and biotechnology,* vol. 82, pp. 1067-78, 2009.

[23] A. Madhavan, *et al.*, "Alcoholic fermentation of xylose and mixed sugars using recombinant Saccharomyces cerevisiae engineered for xylose utilization," *Applied microbiology and biotechnology,* vol. 82, pp. 1037-47, 2009.

[24] D. Brat, *et al.*, "Functional expression of a bacterial xylose isomerase in Saccharomyces cerevisiae," *Applied and environmental microbiology,* vol. 75, pp. 2304-11, 2009.

[25] B. Hahn-Hagerdal, *et al.*, "Metabolic engineering of Saccharomyces cerevisiae for xylose utilization," *Adv Biochem Eng Biotechnol,* vol. 73, pp. 53-84, 2001.

[26] T. W. Jeffries and Y. S. Jin, "Ethanol and thermotolerance in the bioconversion of xylose by yeasts," *Advances in applied microbiology,* vol. 47, pp. 221-68, 2000.

[27] Y.-S. Jin, *et al.*, "Conversion of xylose to ethanol by recombinant saccharomyces cerevisiae containing genes for xylose reductase and xylitol dehydrogenase from pichia stipitis," *J. Microbiol. Biotechnol.,* vol. 10, pp. 564-567, 2000.

[28] M. Sonderegger, *et al.*, "Molecular Basis for Anaerobic Growth of Saccharomyces cerevisiae on Xylose , Investigated by Global Gene Expression and Metabolic Flux Analysis," *APPLIED AND ENVIRONMENTAL MICROBIOLOGY,* vol. 70, pp. 2307-2317, 2004.

[29] Y.-s. Jin, *et al.*, "Saccharomyces cerevisiae Engineered for Xylose Metabolism Exhibits a Respiratory Response," *APPLIED AND ENVIRONMENTAL MICROBIOLOGY,* vol. 70, pp. 6816-6825, 2004.

[30] M. H. Toivari, *et al.*, "Conversion of xylose to ethanol by recombinant Saccharomyces cerevisiae: importance of xylulokinase (XKS1) and oxygen availability," *Metab Eng,* vol. 3, pp. 236-49, Jul 2001.

[31] S. Watanabe, *et al.*, "Complete reversal of coenzyme specificity of xylitol dehydrogenase and increase of thermostability by the introduction of structural zinc," *J Biol Chem,* vol. 280, pp. 10340-9, Mar 18 2005.

[32] L. Liang, *et al.*, "Altering coenzyme specificity of Pichia stipitis xylose reductase by the semi-rational approach CASTing," *Microbial cell factories,* vol. 6, pp. 36-36, 2007.

[33] S. Watanabe, *et al.*, "Ethanol production from xylose by recombinant Saccharomyces cerevisiae expressing protein engineered NADP+-dependent xylitol dehydrogenase," *Journal of biotechnology,* vol. 130, pp. 316-319, 2007.

[34] S. Watanabe, *et al.*, "Ethanol production from xylose by recombinant Saccharomyces cerevisiae expressing protein-engineered NADH-preferring xylose reductase from Pichia stipitis," *Microbiology (Reading, England),* vol. 153, pp. 3044-54, 2007.

[35] S. Watanabe, *et al.*, "The Positive Effect of the Decreased NADPH-Preferring Activity of Xylose Reductase from Pichia stipitis on Ethanol Production Using Xylose-Fermenting Recombinant Saccharomyces cerevisiae," *Bioscience, Biotechnology, and Biochemistry,* vol. 71, pp. 1365-1369, 2007.

[36] J. H. Van Vleet and T. W. Jeffries, "Yeast metabolic engineering for hemicellulosic ethanol production," *Curr Opin Biotechnol,* vol. 20, pp. 300-6, Jun 2009.

[37] B. Petschacher and B. Nidetzky, "Altering the coenzyme preference of xylose reductase to favor utilization of NADH enhances ethanol yield from xylose in a metabolically engineered strain of Saccharomyces cerevisiae," *Microbial cell factories,* vol. 7, pp. 9-9, 2008.

[38] A. H. Ehrensberger, *et al.*, "Structure-guided engineering of xylitol dehydrogenase cosubstrate specificity," *Structure,* vol. 14, pp. 567-75, Mar 2006.

[39] A. Matsushika, *et al.*, "Bioethanol production from xylose by recombinant Saccharomyces cerevisiae expressing xylose reductase, NADP(+)-dependent xylitol dehydrogenase, and xylulokinase," *Journal of bioscience and bioengineering,* vol. 105, pp. 296-9, 2008.

[40] A. Matsushika, *et al.*, "Expression of protein engineered NADP+-dependent xylitol dehydrogenase increases ethanol production from xylose in recombinant Saccharomyces cerevisiae," *Applied microbiology and biotechnology,* vol. 81, pp. 243-255, 2008.

[41] J. Zhang, *et al.*, "Construction of a recombinant S. cerevisiae expressing a fusion protein and study on the effect of converting xylose and glucose to ethanol," *Applied biochemistry and biotechnology,* vol. 150, pp. 185-192, 2008.

[42] K. Karhumaa, *et al.*, "High activity of xylose reductase and xylitol dehydrogenase improves xylose fermentation by recombinant Saccharomyces cerevisiae," *Applied microbiology and biotechnology,* vol. 73, pp. 1039-46, 2007.

[43] A. Matsushika and S. Sawayama, "Efficient bioethanol production from xylose by recombinant saccharomyces cerevisiae requires high activity of xylose reductase and moderate xylulokinase activity," *Journal of bioscience and bioengineering,* vol. 106, pp. 306-9, 2008.

[44] M. Walfridsson, *et al.*, "Ethanolic fermentation of xylose with Saccharomyces cerevisiae harboring the Thermus thermophilus xylA gene, which expresses an active xylose (glucose) isomerase," *Applied and environmental microbiology,* vol. 62, pp. 4648-51, 1996.

[45] M. Gardonyi and B. Hahn-Hägerdal, "The Streptomyces rubiginosus xylose isomerase is misfolded when expressed in Saccharomyces cerevisiae," *Enzyme and Microbial Technology,* vol. 32, pp. 252-259, 2003.

[46] A. J. A. V. Maris, *et al.*, "Development of Efficient Xylose Fermentation in Saccharomyces cerevisiae : Xylose Isomerase as a Key Component," *Adv Biochem Engin/Biotechnol,* vol. 108, pp. 179-204, 2007.

[47] M. Kuyper, *et al.*, "High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by ?," *FEMS Yeast Research,* vol. 4, pp. 69-78, 2003.

[48] J. R. Almeida, *et al.*, "Pichia stipitis xylose reductase helps detoxifying lignocellulosic hydrolysate by reducing 5-hydroxymethyl-furfural (HMF)," *Biotechnol Biofuels,* vol. 1, p. 12, 2008.

[49] M. Walfridsson, *et al.*, "Xylose-metabolizing Saccharomyces cerevisiae strains overexpressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase," *Applied and environmental microbiology,* vol. 61, pp. 4184-90, 1995.

[50] Y.-s. Jin, *et al.*, "Molecular Cloning of XYL3 (D -Xylulokinase) from Pichia stipitis and Characterization of Its Physiological Function," *APPLIED AND ENVIRONMENTAL MICROBIOLOGY,* vol. 68, pp. 1232-1239, 2002.

[51] X. X. Deng and N. W. Ho, "Xylulokinase activity in various yeasts including Saccharomyces cerevisiae containing the cloned xylulokinase gene. Scientific note," *Appl Biochem Biotechnol,* vol. 24-25, pp. 193-9, Spring-Summer 1990.

[52] J. M. Rodriguez-Pena, *et al.*, "The YGR194c (XKS1) gene encodes the xylulokinase from the budding yeast Saccharomyces cerevisiae," *FEMS Microbiol Lett,* vol. 162, pp. 155-60, May 1 1998.

[53] A. Eliasson, *et al.*, "Xylulose fermentation by mutant and wild-type strains of Zygosaccharomyces and Saccharomyces cerevisiae," *Appl Microbiol Biotechnol,* vol. 53, pp. 376-82, Apr 2000.

[54] A. Eliasson, *et al.*, "Anaerobic Xylose Fermentation by Recombinant Saccharomyces Mineral Medium Chemostat Cultures," *APPLIED AND ENVIRONMENTAL MICROBIOLOGY,* vol. 66, pp. 3381-3386, 2000.

[55] B. Teusink, *et al.*, "The danger of metabolic pathways with turbo design," *Trends Biochem Sci,* vol. 23, pp. 162-9, May 1998.

[56] B. Johansson, *et al.*, "Xylulokinase overexpression in two strains of Saccharomyces cerevisiae also expressing xylose reductase and xylitol dehydrogenase and its effect on fermentation of xylose and lignocellulosic hydrolysate," *Appl Environ Microbiol,* vol. 67, pp. 4249-55, Sep 2001.

[57] O. Akinterinwa and P. C. Cirino, "Heterologous expression of D-xylulokinase from Pichia stipitis enables high levels of xylitol production by engineered Escherichia coli growing on xylose," *Metab Eng,* vol. 11, pp. 48-55, Jan 2009.

[58] J. M. Gancedo and C. Gancedo, "Concentrations of intermediary metabolites in yeast," *Biochimie,* vol. 55, pp. 205-11, 1973.

[59] B. Johansson and B. Hahn-Hagerdal, "Overproduction of pentose phosphate pathway enzymes using a new CRE-loxP expression vector for repeated genomic integration in Saccharomyces cerevisiae," *Yeast,* vol. 19, pp. 225-31, Feb 2002.

[60] M. Jeppsson, *et al.*, "Reduced oxidative pentose phosphate pathway flux in recombinant xylose-utilizing Saccharomyces cerevisiae strains improves the ethanol yield from xylose," *Appl Environ Microbiol,* vol. 68, pp. 1604-9, Apr 2002.

[61] B. Johansson and B. Hahn-Hagerdal, "The non-oxidative pentose phosphate pathway controls the fermentation rate of xylulose but not of xylose in Saccharomyces cerevisiae TMB3001," *FEMS Yeast Res,* vol. 2, pp. 277-82, Aug 2002.

[62] K. Karhumaa, *et al.*, "Investigation of limiting metabolic steps in the utilization of xylose by recombinant Saccharomyces cerevisiae using metabolic engineering," *Yeast,* vol. 22, pp. 359-68, 2005.

[63] A. Kuhn, *et al.*, "Purification and partial characterization of an aldo-keto reductase from Saccharomyces cerevisiae," *Appl Environ Microbiol,* vol. 61, pp. 1580-5, Apr 1995.

[64] J. W. Wenger, *et al.*, "Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from Saccharomyces cerevisiae," *PLoS Genet,* vol. 6, p. e1000942, 2010.

[65] K. Yamanaka, "Inhibition of D-xylose isomerase by pentitols and D-lyxose," *Arch Biochem Biophys,* vol. 131, pp. 502-6, May 1969.

[66] K. L. Traff, *et al.*, "Deletion of the GRE3 aldose reductase gene and its influence on xylose metabolism in recombinant strains of Saccharomyces cerevisiae expressing the xylA and XKS1 genes," *Appl Environ Microbiol,* vol. 67, pp. 5668-74, Dec 2001.

[67] A. Lonn, *et al.*, "Xylose isomerase. activity influences xylose fermentation with recombinant Saccharomyces cerevisiae strains expressing mutated xylA from Thermus thermophilus," *Enzyme and Microbial Technology,* vol. 32, pp. 567-573, Apr 8 2003.

[68] M. Koffas, "Evolutionary metabolic engineering," *Metabolic Engineering,* vol. 7, pp. 1-3, Jan 2005.

[69] U. Sauer, "Evolutionary engineering of industrially important microbial phenotypes," *Adv Biochem Eng Biotechnol,* vol. 73, pp. 129-69, 2001.

[70] C. Wahlbom, *et al.*, "Generation of the improved recombinant xylose-utilizing TMB 3400 by random mutagenesis and physiological comparison with CBS 6054," *FEMS Yeast Research,* vol. 3, pp. 319-326, 2003.

[71] M. Sonderegger and U. Sauer, "Evolutionary engineering of Saccharomyces cerevisiae for anaerobic growth on xylose," *Appl Environ Microbiol,* vol. 69, pp. 1990-8, Apr 2003.

[72] J. Becker and E. Boles, "A Modified Saccharomyces cerevisiae Strain That Consumes L - Arabinose and Produces Ethanol," *APPLIED AND ENVIRONMENTAL MICROBIOLOGY,* vol. 69, pp. 4144-4150, 2003.

[73] H. W. Wisselink, *et al.*, "Engineering of Saccharomyces cerevisiae for efficient anaerobic alcoholic fermentation of L-arabinose," *Applied and environmental microbiology,* vol. 73, pp. 4881-91, 2007.

[74] Rosa, *et al.*, "Improved xylose and arabinose utilization by an industrial recombinant Saccharomyces cerevisiae strain using evolutionary engineering," *Biotechnology for Biofuels,* vol. 3, pp. 13-13, 2010.

[75] H. W. Wisselink, *et al.*, "Novel evolutionary engineering approach for accelerated utilization of glucose, xylose, and arabinose mixtures by engineered Saccharomyces cerevisiae strains," *Appl Environ Microbiol,* vol. 75, pp. 907-14, Feb 2009.

[76] C. F. Wahlbom, *et al.*, "Molecular Analysis of a Saccharomyces cerevisiae Mutant with Improved Ability To Utilize Xylose Shows Enhanced Expression of Proteins Involved in Transport , Initial Xylose Metabolism , and the Pentose Phosphate Pathway," *APPLIED AND ENVIRONMENTAL MICROBIOLOGY,* vol. 69, pp. 740-746, 2003.

[77] J.-P. Pitkänen, *et al.*, "Xylose chemostat isolates of Saccharomyces cerevisiae show altered metabolite and enzyme levels compared with xylose, glucose, and ethanol metabolism of the original strain," *Applied microbiology and biotechnology,* vol. 67, pp. 827-37, 2005.

[78] A. Maier, *et al.*, "Characterisation of glucose transport in Saccharomyces cerevisiae with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters," *FEMS yeast research,* vol. 2, pp. 539-50, 2002.

[79] M. Sedlak and N. W. Y. Ho, "Characterization of the effectiveness of hexose transporters for transporting xylose during glucose and xylose co-fermentation by a recombinant Saccharomyces yeast," *Yeast,* vol. 21, pp. 671-84, 2004.

[80] M. J. Leandro, *et al.*, "Hexose and pentose transport in ascomycetous yeasts: an overview," *FEMS Yeast Res,* vol. 9, pp. 511-25, Jun 2009.

[81] A. L. Kruckeberg, "The hexose transporter family of Saccharomyces cerevisiae," *Arch Microbiol,* vol. 166, pp. 283-92, Nov 1996.

[82] S. Ozcan and M. Johnston, "Function and regulation of yeast hexose transporters," *Microbiol Mol Biol Rev,* vol. 63, pp. 554-69, Sep 1999.

[83] M. Silverman, "Structure and function of hexose transporters," *Annual review of biochemistry,* vol. 60, pp. 757-94, 1991.

[84] M. J. Leandro, *et al.*, "Two glucose/xylose transporter genes from the yeast Candida intermedia: first molecular characterization of a yeast xylose-H+ symporter," *The Biochemical journal,* vol. 395, pp. 543-9, 2006.

[85] T. Hamacher, *et al.*, "Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization," *Microbiology,* vol. 148, pp. 2783-8, 2002.

[86] P. Kötter and M. Ciriacy, "Xylose fermentation by Saccharomyces cerevisiae," *Applied Microbiology and Biotechnology,* vol. 38, pp. 776-783, 1993.

[87] W. J. Lee, *et al.*, "Kinetic studies on glucose and xylose transport in Saccharomyces cerevisiae," *Applied microbiology and biotechnology,* vol. 60, pp. 186-91, 2002.

[88] M. Gárdonyi, *et al.*, "Control of xylose consumption by xylose transport in recombinant Saccharomyces cerevisiae," *Biotechnology and bioengineering,* vol. 82, pp. 818-24, 2003.

[89] A. Saloheimo, *et al.*, "Xylose transport studies with xylose-utilizing Saccharomyces cerevisiae strains expressing heterologous and homologous permeases," *Applied microbiology and biotechnology,* vol. 74, pp. 1041-52, 2007.

[90] a. Nobre, *et al.*, "Transport and utilization of hexoses and pentoses in the halotolerant yeast Debaryomyces hansenii," *Applied and environmental microbiology,* vol. 65, pp. 3594-8, 1999.

[91] T. Weierstall, *et al.*, "Cloning and characterization of three genes (SUT1-3) encoding glucose transporters of the yeast Pichia stipitis," *Molecular microbiology,* vol. 31, pp. 871-83, 1999.

[92] B. U. Stambuk, *et al.*, "D-Xylose Transport by Candida succiphila and Kluyveromyces marxianus," *Applied Biochemistry And Biotechnology,* vol. 105, pp. 255-263, 2003.

[93] M. J. Leandro, *et al.*, "The expression in Saccharomyces cerevisiae of a glucose/xylose symporter from Candida intermedia is affected by the presence of a glucose/xylose facilitator," *Microbiology (Reading, England),* vol. 154, pp. 1646-55, 2008.

[94] D. Runquist, *et al.*, "Expression of the Gxf1 transporter from Candida intermedia improves fermentation performance in recombinant xylose-utilizing Saccharomyces cerevisiae," *Applied microbiology and biotechnology,* vol. 82, pp. 123-30, 2009.

[95] T. Jojima, *et al.*, "Sugar transporters in efficient utilization of mixed sugar substrates: current knowledge and outlook," *Appl Microbiol Biotechnol,* vol. 85, pp. 471-80, Jan 2010.

[96] M. Buttner and N. Sauer, "Monosaccharide transporters in plants: structure, function and physiology," *Biochim Biophys Acta,* vol. 1465, pp. 263-74, May 1 2000.

[97] H. Sun, *et al.*, "Double-Strand Breaks at an Initiation Site for Meiotic Gene Conversion," *Nature,* vol. 338, pp. 87-90, Mar 2 1989.

[98] R. S. Sikorski and P. Hieter, "A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae," *Genetics,* vol. 122, pp. 19-27, May 1989.

[99] D. Mumberg, *et al.*, "Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds," *Gene,* vol. 156, pp. 119-22, Apr 14 1995.

[100] J. Sambrook and D. Russell, *Molecular Cloning: A Laboratory Manual (Third Edition)*: Cold Spring Harbor Laboratory Press, 2001.

[101] M. Bettiga, *et al.*, "Comparing the xylose reductase/xylitol dehydrogenase and xylose isomerase pathways in arabinose and xylose fermenting Saccharomyces cerevisiae strains," *Biotechnology for biofuels,* vol. 1, pp. 16-16, 2008.

[102] S. Coyle and E. Kroll, "Starvation induces genomic rearrangements and starvation-resilient phenotypes in yeast," *Mol Biol Evol,* vol. 25, pp. 310-8, Feb 2008.

[103] D. A. Thompson, *et al.*, "Ploidy controls the success of mutators and nature of mutations during budding yeast evolution," *Curr Biol,* vol. 16, pp. 1581-90, Aug 22 2006.

[104] P. L. Foster, "Stress-induced mutagenesis in bacteria," *Crit Rev Biochem Mol Biol,* vol. 42, pp. 373-97, Sep-Oct 2007.

[105] J. E. Bailey, *et al.*, "Inverse metabolic engineering: A strategy for directed genetic engineering of useful phenotypes," *Biotechnology and Bioengineering,* vol. 52, pp. 109-121, Oct 5 1996.

[106] J. E. Bailey, *et al.*, "Inverse metabolic engineering: A strategy for directed genetic engineering of useful phenotypes," *Biotechnology and Bioengineering,* vol. 79, pp. 568-579, Sep 5 2002.

[107] B. L. Wang, "High throughput screen for cells with high extracellular metabolite consumption--secretion rates using microfluidic droplets," Ph. D., Dept. of Chemical Engineering., Massachusetts Institute of Technology, Cambridge, MA, 2009.

[108] H. Ni, *et al.*, "Transposon mutagenesis to improve the growth of recombinant Saccharomyces cerevisiae on D-xylose," *Applied and environmental microbiology,* vol. 73, pp. 2061-2066, 2007.

[109] G. F. Carle and M. V. Olson, "An electrophoretic karyotype for yeast," *Proc Natl Acad Sci U S A,* vol. 82, pp. 3756-60, Jun 1985.

[110] H. Kerstershilderson, *et al.*, "Kinetic Characterization of D-Xylose Isomerases by Enzymatic Assays Using D-Sorbitol Dehydrogenase," *Enzyme and Microbial Technology,* vol. 9, pp. 145-148, Mar 1987.

[111] D. K. Shamanna and K. E. Sanderson, "Uptake and Catabolism of D-Xylose in Salmonella-Typhimurium Lt2," *JOURNAL OF BACTERIOLOGY,* vol. 139, pp. 64-70, 1979.

[112] J. Hauf, *et al.*, "Simultaneous genomic overexpression of seven glycolytic enzymes in the yeast Saccharomyces cerevisiae," *Enzyme Microb Technol,* vol. 26, pp. 688-698, Jun 1 2000.

[113] H. Alper, *et al.*, "Tuning genetic control through promoter engineering," *PNAS,* vol. 102, pp. 12678-12683, 2005.

[114] K. L. Traff, *et al.*, "Putative xylose and arabinose reductases in Saccharomyces cerevisiae," *Yeast,* vol. 19, pp. 1233-41, Oct 2002.

[115] J. T. Pronk, "Auxotrophic yeast strains in fundamental and applied research," *Appl Environ Microbiol,* vol. 68, pp. 2095-100, May 2002.

[116] R. J. Baerends, *et al.*, "Impaired uptake and/or utilization of leucine by Saccharomyces cerevisiae is suppressed by the SPT15-300 allele of the TATA-binding protein gene," *Appl Environ Microbiol,* vol. 75, pp. 6055-61, Oct 2009.

[117] A. A. Andreasen and T. J. Stier, "Anaerobic nutrition of Saccharomyces cerevisiae. I. Ergosterol requirement for growth in a defined medium," *J Cell Physiol,* vol. 41, pp. 23-36, Feb 1953.

[118] A. A. Andreasen and T. J. Stier, "Anaerobic nutrition of Saccharomyces cerevisiae. II. Unsaturated fatty acid requirement for growth in a defined medium," *J Cell Physiol,* vol. 43, pp. 271-81, Jun 1954.

[119] K. Ohta and S. Hayashida, "Role of tween 80 and monoolein in a lipid-sterol-protein complex which enhances ethanol tolerance of sake yeasts," *Appl Environ Microbiol,* vol. 46, pp. 821-5, Oct 1983.

[120] C. Bro, *et al.*, "In silico aided metabolic engineering of Saccharomyces cerevisiae for improved bioethanol production," *Metab Eng,* vol. 8, pp. 102-11, Mar 2006.

Appendix

A.1. Heterologous xylose isomerases

Piromyces sp. E2 xylose isomerase

PROTEIN LENGTH = 437

MAKEYFPQIQKIKFEGKDSKNPLAFHYYDAEKEVMGKKMKDWLRFAMAWWHTLCAEGADQFGGGTKSFPWNE
GTDAIEIAKQKVDAGFEIMQKLGIPYYCFHDVDLVSEGNSIEEYESNLKAVVAYLKEKQKETGIKLLWSTAN
VFGHKRYMNGASTNPDFDVVARAIVQIKNAIDAGIELGAENYVFWGGREGYMSLLNTDQKREKEHMATMLTM
ARDYARSKGFKGTFLIEPKPMEPTKHQYDVDTETAIGFLKAHNLDKDFKVNIEVNHATLAGHTFEHELACAV
DAGMLGSIDANRGDYQNGWDTDQFPIDQYELVQAWMEIIRGGGFVTGGTNFDAKTRRNSTDLEDIIIAHVSG
MDAMARALENAAKLLQESPYTKMKKERYASFDSGIGKDFEDGKLTLEQVYEYGKKNGEPKQTSGKQELYEAI
VAMYQ*

DNA LENGTH = 1314 (codon optimized)

ATGGCTAAAGAGTACTTCCCACAGATTCAGAAGATAAAGTTCGAGGGCAAAGATTCTAAAAACCCTTTGGCT
TTCCACTACTATGATGCAGAGAAGGAAGTCATGGGAAAGAAAATGAAGGATTGGTTGAGATTTGCTATGGCT
TGGTGGCATACTTTGTGTGCTGAAGGTGCAGACCAGTTCGGCGGTGGCACTAAGTCTTTTCCTTGGAATGAG
GGTACTGATGCCATTGAAATCGCCAAACAAAAGGTAGACGCTGGTTTTGAGATCATGCAGAAGTTGGGCATC
CCTTATTACTGTTTTCACGATGTCGATTTGGTGAGTGAAGGCAATAGTATAGAGGAATACGAGTCTAACTTA
AAGGCAGTCGTTGCCTATTTGAAGGAGAAGCAAAAGGAAACTGGTATCAAATTGTTGTGGAGTACTGCTAAC
GTCTTCGGCCACAAAAGATACATGAACGGTGCTTCTACTAATCCAGACTTTGATGTAGTCGCTAGAGCTATA
GTCCAGATTAAGAATGCTATCGACGCCGGAATTGAGTTGGGAGCTGAGAACTATGTTTTTTGGGGAGGTAGG
GAAGGCTATATGTCTTTGTTGAATACTGACCAGAAGAGAGAGAAAGAACACATGGCAACAATGTTAACTATG
GCAAGAGATTACGCAAGGAGTAAGGGCTTTAAGGGCACTTTTTTGATTGAACCTAAGCCTATGGAACCAACT
AAACACCAATATGATGTTGACACTGAAACAGCCATCGGTTTCTTGAAGGCCCACAACTTGGATAAAGATTTT
AAGGTAAACATTGAGGTCAATCACGCCACCTTGGCCGGTCACACTTTCGAACATGAATTGGCTTGTGCTGTT
GATGCTGGAATGTTGGGTTCTATTGATGCAAATAGAGGCGATTATCAGAATGGTTGGGATACTGATCAATTT
CCAATCGACCAATACGAATTGGTTCAAGCCTGGATGGAAATCATAAGAGGTGGTGGCTTTGTAACTGGTGGA
ACTAACTTCGATGCCAAAACAAGAAGAAACTCCACTGACTTGGAGGATATCATTATTGCTCACGTTTCCGGT
ATGGATGCAATGGCCAGGGCCTTGGAGAACGCTGCTAAGTTGTTACAAGAATCCCCCTACACTAAGATGAAG
AAAGAGAGGTACGCATCATTCGATTCTGGAATCGGCAAGGATTTTGAGGACGGAAAGTTGACTTTAGAGCAG
GTTTATGAGTACGGTAAAAAGAATGGCGAGCCTAAACAAACCTCTGGTAAGCAGGAATTGTACGAAGCTATT
GTCGCAATGTATCAATAA

50 Oligonucleotides synthesized

---

1 ATGGCTAAAGAGTACTTCCCACAG 24
2 CTTTGCCCTCGAACTTTATCTTCTGAATCTGTGGGAAGTACTCTTTAGCCA 51
3 CAGAAGATAAAGTTCGAGGGCAAAGATTCTAAAAACCCTTTGGCTTTCCAC 51
4 ATGACTTCCTTCTCTGCATCATAGTAGTGGAAAGCCAAAGGGTTTTTAGAA 51
5 ACTATGATGCAGAGAAGGAAGTCATGGGAAAGAAAATGAAGGATTGGTTGA 51
6 GTATGCCACCAAGCCATAGCAAATCTCAACCAATCCTTCATTTTCTTTCCC 51
7 TGCTATGGCTTGGTGGCATACTTTGTGTGCTGAAGGTGCAGACCAGTTCGG 51
8 CTCATTCCAAGGAAAAGACTTAGTGCCACCGCCGAACTGGTCTGCACCTTC 51
9 GCACTAAGTCTTTTCCTTGGAATGAGGGTACTGATGCCATTGAAATCGCCA 51
10 TGATCTCAAAACCAGCGTCTACCTTTTGTTTGGCGATTTCAATGGCATCAG 51
11 GGTAGACGCTGGTTTTGAGATCATGCAGAAGTTGGGCATCCCTTATTACTG 51
12 TCACTCACCAAATCGACATCGTGAAAACAGTAATAAGGGATGCCCAACTTC 51
13 ACGATGTCGATTTGGTGAGTGAAGGCAATAGTATAGAGGAATACGAGTCTA 51
14 AGGCAACGACTGCCTTTAAGTTAGACTCGTATTCCTCTATACTATTGCC 49
15 ACTTAAAGGCAGTCGTTGCCTATTTGAAGGAGAAGCAAAAGGAAACTGGTA 51
16 GTTAGCAGTACTCCACAACAATTTGATACCAGTTTCCTTTTGCTTCTCCTT 51
17 TCAAATTGTTGTGGAGTACTGCTAACGTCTTCGGCCACAAAAGATACATGA 51
18 CAAAGTCTGGATTAGTAGAAGCACCGTTCATGTATCTTTTGTGGCCGAAGA 51
19 GGTGCTTCTACTAATCCAGACTTTGATGTAGTCGCTAGAGCTATAGTCCAG 51
20 AATTCCGGCGTCGATAGCATTCTTAATCTGGACTATAGCTCTAGCGACTAC 51
21 AATGCTATCGACGCCGGAATTGAGTTGGGAGCTGAGAACTATGTTTTTTGG 51
22 AAAGACATATAGCCTTCCCTACCTCCCCAAAAAACATAGTTCTCAGCTCCC 51
23 GAGGTAGGGAAGGCTATATGTCTTTGTTGAATACTGACCAGAAGAGAGAGA 51
24 GTTAACATTGTTGCCATGTGTTCTTTCTCTCTCTTCTGGTCAGTATTCAAC 51
25 AAGAACACATGGCAACAATGTTAACTATGGCAAGAGATTACGCAAGGAGTA 51
26 TCAATCAAAAAAGTGCCCTTAAAGCCCTTACTCCTTGCGTAATCTCTTGCC 51
27 GCTTTAAGGGCACTTTTTTGATTGAACCTAAGCCTATGGAACCAACTAAAC 51
28 TGTTTCAGTGTCAACATCATATTGGTGTTTAGTTGGTTCCATAGGCTTAGG 51
29 ACCAATATGATGTTGACACTGAAACAGCCATCGGTTTCTTGAAGGCCCACA 51
30 TCAATGTTTACCTTAAAATCTTTATCCAAGTTGTGGGCCTTCAAGAAACCG 51
31 ACTTGGATAAAGATTTTAAGGTAAACATTGAGGTCAATCACGCCACCTTGG 51
32 CAAGCCAATTCATGTTCGAAAGTGTGACCGGCCAAGGTGGCGTGATTGACC 51
33 CACTTTCGAACATGAATTGGCTTGTGCTGTTGATGCTGGAATGTTGGGTTC 51
34 CTGATAATCGCCTCTATTTGCATCAATAGAACCCAACATTCCAGCATCAAC 51
35 TGATGCAAATAGAGGCGATTATCAGAATGGTTGGGATACTGATCAATTTCC 51
36 CTTGAACCAATTCGTATTGGTCGATTGGAAATTGATCAGTATCCCAACCAT 51
37 ATCGACCAATACGAATTGGTTCAAGCCTGGATGGAAATCATAAGAGGTGGT 51
38 AAGTTAGTTCCACCAGTTACAAAGCCACCACCTCTTATGATTTCCATCCAG 51
39 GCTTTGTAACTGGTGGAACTAACTTCGATGCCAAAACAAGAAGAAACTCCA 51

40 AGCAATAATGATATCCTCCAAGTCAGTGGAGTTTCTTCTTGTTTTGGCATC 51
41 CTGACTTGGAGGATATCATTATTGCTCACGTTTCCGGTATGGATGCAATGG 51
42 CAACTTAGCAGCGTTCTCCAAGGCCCTGGCCATTGCATCCATACCGGAAAC 51
43 CTTGGAGAACGCTGCTAAGTTGTTACAAGAATCCCCCTACACTAAGATGAA 51
44 AATCGAATGATGCGTACCTCTCTTTCTTCATCTTAGTGTAGGGGGATTCTT 51
45 AAAGAGAGGTACGCATCATTCGATTCTGGAATCGGCAAGGATTTTGAGGAC 51
46 CATAAACCTGCTCTAAAGTCAACTTTCCGTCCTCAAAATCCTTGCCGATTC 51
47 GAAAGTTGACTTTAGAGCAGGTTTATGAGTACGGTAAAAAGAATGGCGAGC 51
48 AATTCCTGCTTACCAGAGGTTTGTTTAGGCTCGCCATTCTTTTTACCGTAC 51
49 ACAAACCTCTGGTAAGCAGGAATTGTACGAAGCTATTGTCGCAATGTATCA 51
50 TTGATACATTGCGACAATAGCTTCG 25

Thermus thermophilus xylose isomerase

PROTEIN LENGTH = 387

MYEPKPEHRFTFGLWTVGNVGRDPFGDAVRERLDPVYVVHKLAELGAYGVNLHDEDLIPRGTPPQERDQIVR
RFKKALDETGLKVPMVTANLFSDPAFKDGAFTSPDPWVRAYALRKSLETMDLGAELGAEIYVVWPGREGAEV
EATGKARKVWDWVREALNFMAAYAEDQGYGYRFALEPKPNEPRGDIYFATVGSMLAFIHTLDRPERFGLNPE
FAHETMAGLNFVHAVAQALDAGKLFHIDLNDQRMSRFDQDLRFGSENLKAAFFLVDLLESSGYQGPRHFDAH
ALRTEDEEGVWAFARGCMRTYLILKERAEAFREDPEVKELLAAYYQEDPAALALLGPYSREKAEALKRAELP
LEAKRRRGYALERLDQLAVEYLLGVRG*

DNA LENGTH = 1164 (Expression optimized)

ATGTATGAACCCAAACCCGAACATAGATTTACATTTGGACTTTGGACAGTAGGAAATGTTGGTAGAGATCCA
TTCGGTGACGCAGTTAGAGAGAGATTAGATCCAGTTTATGTTGTTCACAAATTGGCTGAATTAGGCGCATAT
GGTGTTAATTTACATGATGAAGATTTAATTCCCAGAGGTACACCCCCACAAGAAAGAGACCAAATTGTAAGA
AGATTTAAAAAAGCACTAGACGAAACAGGACTAAAAGTACCAATGGTAACAGCAAATTTATTTTCAGACCCC
GCTTTCAAAGATGGAGCATTCACTTCTCCCGATCCGTGGGTTAGAGCTTATGCCTTGAGAAAAAGTTTAGAA
ACAATGGATTTGGGTGCAGAATTAGGAGCAGAAATATACGTTGTATGGCCAGGAAGAGAAGGAGCAGAAGTA
GAGGCAACAGGAAAAGCAAGAAAGGTGTGGGATTGGGTAAGGGAAGCCCTAAATTTTATGGCTGCTTATGCT
GAAGATCAAGGATACGGATATAGATTCGCACTAGAACCAAAACCGAACGAGCCAAGAGGTGATATATATTTT
GCTACAGTCGGTTCTATGTTGGCATTTATACATACGTTAGACAGGCCTGAAAGATTCGGATTGAACCCAGAA
TTTGCACACGAAACAATGGCCGGATTAAACTTTGTACACGCTGTTGCTCAAGCCTTAGACGCTGGTAAATTA
TTTCACATAGACCTAAATGACCAAAGAATGTCACGTTTCGACCAAGATTTAAGATTTGGTTCAGAGAATTTA
AAAGCTGCATTTTTCCTAGTAGATTTATTGGAATCAAGTGGTTATCAAGGACCAAGACATTTTGACGCACAC
GCACTAAGAACTGAAGATGAAGAAGGGGTTTGGGCCTTTGCCAGAGGTTGCATGAGAACGTATTTAATACTT
AAAGAAAGAGCCGAAGCATTTAGAGAAGATCCAGAAGTTAAAGAATTGTTAGCTGCCTACTACCAAGAAGAC
CCAGCTGCACTAGCATTACTAGGACCGTATTCAAGAGAAAAAGCTGAAGCATTAAAAAGAGCCGAATTACCA
TTAGAAGCAAAGAGAAGAAGAGGTTACGCATTAGAAAGATTAGACCAATTAGCTGTAGAATATTTATTAGGA
GTCAGGGGATAA

E.coli xylose isomerase

PROTEIN LENGTH = 440

MQAYFDQLDRVRYEGSKSSNPLAFRHYNPDELVLGKRMEEHLRFAACYWHTFCWNGADMFGVGAFNRPWQQP
GEALALAKRKADVAFEFFHKLHVPFYCFHDVDVSPEGASLKEYINNFAQMVDVLAGKQEESGVKLLWGTANC
FTNPRYGAGAATNPDPEVFSWAATQVVTAMEATHKLGGENYVLWGGREGYETLLNTDLRQEREQLGRFMQMV
VEHKHKIGFQGTLLIEPKPQEPTKHQYDYDAATVYGFLKQFGLEKEIKLNIEANHATLAGHSFHHEIATAIA
LGLFGSVDANRGDAQLGWDTDQFPNSVEENALVMYEILKAGGFTTGGLNFDAKVRRQSTDKYDLFYGHIGAM
DTMALALKIAARMIEDGELDKRIAQRYSGWNSELGQQILKGQMSLADLAKYAQEHHLSPVHQSGRQEQLENL
VNHYLFDK*

DNA LENGTH = 1323 (Expression optimized)

ATGCAAGCTTATTTTGATCAATTAGACAGAGTCAGATACGAAGGATCAAAATCATCAAACCCATTAGCATTT
AGACATTATAACCCTGATGAATTAGTATTGGGAAAGAGAATGGAAGAACATCTAAGATTTGCAGCTTGCTAT
TGGCACACTTTTTGTTGGAACGGTGCAGATATGTTTGGGGTGGGTGCATTCAATAGACCATGGCAACAACCC
GGAGAAGCTTTAGCTTTGGCAAAGAGAAAAGCCGACGTGGCCTTTGAATTTTTTCACAAACTTCATGTTCCA
TTTTACTGTTTCCATGACGTAGATGTTTCACCCGAAGGTGCATCATTGAAGGAATATATAAACAACTTTGCT
CAGATGGTTGATGTATTAGCTGGTAAACAAGAGGAATCCGGAGTAAAATTATTATGGGGGACAGCGAATTGC
TTTACTAATCCAAGATATGGTGCAGGCGCTGCTACAAATCCAGACCCTGAAGTATTCTCATGGGCAGCTACT
CAAGTCGTTACCGCAATGGAAGCTACGCACAAATTAGGCGGCGAAAATTATGTTTTATGGGGGGGAAGAGAA
GGTTATGAAACTTTATTAAATACAGATTTAAGACAAGAAAGAGAACAATTAGGACGTTTTATGCAAATGGTG
GTAGAACATAAACACAAAATCGGATTTCAAGGAACATTATTAATCGAACCAAAACCACAAGAGCCAACCAAA
CACCAATATGATTACGACGCAGCAACAGTTTACGGATTTTTAAAACAATTTGGGTTAGAAAAGGAAATAAAA
TTGAATATAGAAGCAAATCACGCAACATTAGCAGGACATTCTTTTCATCATGAAATTGCCACCGCTATCGCA
CTTGGTTTATTTGGCAGTGTTGACGCAAATAGAGGAGATGCACAATTAGGATGGGATACTGACCAATTTCCA
AACTCCGTTGAAGAAAATGCATTGGTTATGTATGAGATCTTGAAGGCTGGGGGATTTACAACTGGGGGATTA
AATTTCGATGCTAAAGTTAGAAGGCAATCTACTGATAAATACGATTTATTTTACGGACACATAGGAGCAATG
GATACAATGGCTTTAGCTTTAAAAATCGCTGCAAGAATGATAGAAGATGGCGAATTAGATAAGAGAATAGCA
CAAAGATATTCAGGCTGGAACAGTGAATTAGGACAACAAATTTTAAAAGGGCAAATGTCACTAGCAGACTTA
GCGAAATATGCACAAGAACACCATTTATCACCCGTGCATCAATCTGGTAGACAAGAACAATTAGAAAATTTG
GTAAATCATTATCTATTTGACAAATAA

Bacteroides thetaiotaomicron xylose isomerase

PROTEIN LENGTH = 438

MATKEFFPGIEKIKFEGKDSKNPMAFRYYDAEKVINGKKMKDWLRFAMAWWHTLCAEGGDQFGGGTKQFPWN
GNADAIQAAKDKMDAGFEFMQKMGIEYYCFHDVDLVSEGASVEEYEANLKEIVAYAKQKQAETGIKLLWGTA
NVFGHARYMNGAATNPDFDVVARAAVQIKNAIDATIELGGENYVFWGGREGYMSLLNTDQKREKEHLAQMLT
IARDYARARGFKGTFLIEPKPMEPTKHQYDVDTETVIGFLKAHGLDKDFKVNIEVNHATLAGHTFEHELAVA
VDNGMLGSIDANRGDYQNGWDTDQFPIDNYELTQAMMQIIRNGGLGTGGTNFDAKTRRNSTDLEDIFIAHIA
GMDAMARALESAAALLDESPYKKMLADRYASFDGGKGKEFEDGKLTLEDVVAYAKTKGEPKQTSGKQELYEA
ILNMYC*

DNA LENGTH = 1317 (Expression optimized)

ATGGCTACAAAAGAATTCTTTCCGGGAATCGAAAAAATAAAGTTTGAAGGTAAAGATTCAAAAAACCCAATG
GCTTTTAGATATTATGATGCTGAAAAGGTAATAAACGGAAAAAAAATGAAAGACTGGTTGAGATTTGCTATG
GCTTGGTGGCATACATTATGCGCTGAAGGTGGGGATCAATTCGGGGGTGGGACGAAACAATTCCCTTGGAAC
GGTAATGCAGACGCAATTCAAGCGGCTAAGGATAAAATGGATGCAGGGTTTGAATTTATGCAAAAAATGGGC
ATTGAATACTATTGTTTCCACGACGTTGATTTAGTAAGTGAAGGGGCATCAGTAGAAGAATATGAAGCTAAC
CTTAAAGAAATCGTTGCATACGCAAAACAAAAGCAGGCTGAAACTGGTATAAAATTGTTATGGGGAACAGCT
AACGTTTTCGGACATGCCAGATATATGAATGGTGCAGCTACCAACCCAGATTTTGATGTCGTTGCAAGAGCT
GCTGTACAAATCAAAAATGCGATAGATGCTACAATAGAACTAGGAGGAGAAAATTACGTTTTTTGGGGAGGA
AGAGAAGGGTACATGTCATTACTTAATACGGATCAAAAAAGAGAAAAGGAACACTTAGCCCAAATGTTAACA
ATTGCAAGAGATTACGCAAGAGCAAGAGGATTTAAAGGAACATTTTTAATAGAACCAAAACCGATGGAACCA
ACTAAACATCAATACGACGTAGACACTGAAACAGTTATAGGTTTCCTAAAAGCACACGGACTAGATAAGGAT
TTTAAAGTTAATATAGAAGTGAACCATGCAACTTTAGCTGGACATACATTCGAACATGAGCTAGCGGTAGCA
GTCGATAATGGAATGCTAGGATCTATAGATGCTAATAGAGGAGATTACCAAAACGGCTGGGACACTGATCAA
TTCCCTATTGATAACTATGAATTAACACAAGCAATGATGCAAATAATTAGGAATGGCGGCTTAGGGACAGGA
GGAACAAATTTTGATGCGAAAACTAGAAGAAACTCGACGGATTTAGAAGATATATTTATAGCGCATATTGCT
GGAATGGATGCAATGGCTCGTGCACTGGAATCTGCTGCGGCACTATTAGACGAATCGCCATACAAAAAAATG
CTTGCCGATAGATACGCTTCATTTGACGGGGGGAAAGGAAAAGAATTTGAAGATGGTAAATTGACACTAGAA
GATGTTGTTGCTTACGCTAAGACTAAAGGAGAACCAAAACAAACTTCAGGAAAACAAGAATTATATGAAGCA
ATTTTGAATATGTATTGTTAA

Arabidopsis thaliana xylose isomerase

PROTEIN LENGTH = 477

MKKVEFFMLLLCFIAASSLVSADPPTCPADLGGKCSDSDDWQGDFFPEIPKIKYEGPSSKNPLAYRWYNAEE
EILGKKMKDWFRFSVAFWHTFRGTGGDPFGAATKYWPWEDGTNSVSMAKRRMRANFEFLKKLGVDWWCFHDR
DIAPDGTTLEESNKNLDEVIELAKELQKGSKIKPLWGTAQLFLHPRYMHGGATSSEVGVYAYAAAQVKKAME
VTHYLGGENYVFWGGREGYQTLLNTDMGRELDHLARFFEAAVAYKKKIGFKGTLLIEPKPQEPTKHQYDWDA
ATAANFLRKYGLIDEFKLNIECNHATLSGHTCHHELETARINGLLGNIDANTGDAQTGWDTDQFLTDVGEAT
MVMMSVIKNGGIAPGGFNFDAKLRRESTDVEDLFIAHISGMDTMARGLRNAVKILEEGSLSELVRKRYATWD
SELGKQIEEGKADFEYLEKKAKEFGEPKVSSAKQELAEMIFQSAM*

DNA LENGTH = 1434 (Expression optimized)

ATGAAAAAAGTCGAATTTTTTATGTTATTATTATGTTTTATAGCAGCGTCTAGTTTGGTTTCAGCAGACCCA
CCTACTTGTCCTGCAGATTTAGGAGGCAAATGTTCAGATTCAGATGATTGGCAAGGTGATTTTTTTCCAGAA
ATACCGAAAATTAAGTATGAAGGTCCTTCGAGCAAAAATCCCCTAGCATACAGATGGTATAATGCAGAAGAA
GAGATATTGGGCAAAAAAAATGAAAGATTGGTTTAGGTTTTCAGTCGCTTTCTGGCATACATTTAGAGGCACA
GGTGGTGATCCCTTTGGCGCAGCTACGAAATATTGGCCCTGGGAAGACGGAACAAATAGTGTTAGTATGGCT
AAAAGGAGAATGAGGGCAAATTTTGAATTTTTGAAAAAACTTGGAGTTGATTGGTGGTGTTTTCACGACAGA
GATATCGCACCAGACGGGACAACATTAGAGGAATCGAATAAAAATTTAGATGAAGTTATTGAATTAGCAAAA
GAACTACAAAAAGGGAGTAAAATAAAACCACTATGGGGAACAGCACAATTATTCTTACATCCTCGTTACATG
CATGGAGGTGCAACATCATCAGAAGTCGGAGTTTACGCATATGCAGCCGCACAAGTTAAGAAAGCGATGGAA
GTTACACATTATTTAGGAGGTGAAAATTATGTATTTTGGGGTGGAAGAGAAGGTTATCAAACATTGTTAAAC
ACAGATATGGGCAGAGAATTGGATCATTTAGCAAGATTTTTCGAAGCCGCAGTTGCCTACAAGAAAAAAATC
GGATTTAAAGGAACTTTATTAATTGAACCAAAACCACAAGAGCCAACAAAACATCAATACGATTGGGATGCT
GCCACTGCCGCGAATTTTTTGAGAAAGTATGGGTTAATCGACGAATTTAAATTAAATATAGAATGTAATCAT
GCTACATTGTCCGGCCACACATGCCACCATGAATTAGAAACAGCAAGGATTAATGGATTATTGGGTAACATA
GATGCTAATACAGGTGACGCACAAACAGGATGGGATACAGATCAATTTCTTACAGATGTAGGTGAAGCTACA
ATGGTTATGATGTCAGTTATTAAAAATGGGGGAATCGCACCAGGGGGATTTAATTTCGATGCAAAATTAAGA
AGGGAAAGTACAGACGTAGAGGATTTATTTATTGCACACATATCAGGTATGGATACTATGGCTAGAGGTTTA
AGAAATGCAGTCAAGATATTGGAAGAAGGCTCACTATCAGAATTAGTTAGAAAAAGGTACGCAACATGGGAT
TCGGAATTAGGAAAACAGATAGAAGAAGGTAAAGCAGATTTTGAATACTTAGAAAAAAAGGCAAAGGAATTC
GGTGAACCTAAAGTAAGCTCAGCAAAACAAGAATTGGCCGAAATGATATTTCAAAGTGCTATGTAA

Bacillus subtilis xylose isomerase

PROTEIN LENGTH = 440

MAQSHSSSVNYFGSVNKVVFEGKASTNPLAFKYYNPQEVIGGKTMKEHLRFSIAYWHTFTADGTDVFGAATM
QRPWDHYKGMDLARARVEAAFEMFEKLDAPFFAFHDRDIAPEGSTLKETNQNLDIIVGMIKDYMRDSNVKLL
WNTANMFTNPRFVHGAATSCNADVFAYAAAQVKKGLETAKELGAENYVFWGGREGYETLLNTDLKFELDNLA
RFMHMAVDYAKEIEYTGQFLIEPKPKEPTTHQYDTDAATTIAFLKQYGLDNHFKLNLEANHATLAGHTFEHE
LRMARVHGLLGSVDANQGHPLLGWDTDEFPTDLYSTTLAMYEILQNGGLGSGGLNFDAKVRRSSFEPDDLVY
AHIAGMDAFARGLKVAHKLIEDRVFEDVIQHRYRSFTEGIGLEITEGRANFHTLEQYALNNKTIKNESGRQE
RLKPILNQ*

DNA LENGTH = 1323 (Expression optimized)

ATGGCTCAATCTCATTCTAGTTCAGTTAACTATTTTGGAAGCGTAAACAAAGTGGTTTTCGAAGGGAAAGCT
TCCACTAATCCTTTAGCATTTAAATATTATAATCCTCAAGAAGTAATCGGCGGAAAAACGATGAAAGAGCAT
TTGCGATTTTCTATTGCCTATTGGCATACATTTACTGCTGATGGCACAGACGTTTTTGGAGCAGCTACAATG
CAAAGACCATGGGATCACTATAAAGGCATGGATCTAGCTAGGGCAAGAGTAGAAGCAGCATTTGAGATGTTT
GAAAAACTAGATGCACCATTTTTTGCTTTTCATGATCGAGATATTGCACCAGAAGGAAGTACGTTAAAAGAG
ACAAATCAAAATTTAGATATTATCGTGGGCATGATTAAGGATTACATGAGAGATAGCAACGTTAAGTTATTA
TGGAATACTGCAAACATGTTTACGAACCCCCGTTTCGTCCATGGAGCCGCGACTTCTTGTAATGCAGATGTG
TTTGCGTATGCTGCAGCACAAGTAAAAAAAGGGTTAGAAACAGCAAAAGAGCTTGGCGCCGAGAACTATGTA
TTTTGGGGCGGCCGTGAAGGATACGAAACATTGTTAAATACCGATTTAAAATTTGAGCTTGATAATTTGGCG
AGATTTATGCATATGGCAGTAGATTATGCGAAGGAAATCGAGTATACAGGGCAGTTTTTGATTGAACCAAAA
CCAAAAGAGCCGACCACCCATCAATATGATACAGATGCAGCAACAACCATTGCCTTTTTGAAGCAATATGGC
TTAGACAATCATTTTAAATTAAATCTAGAAGCCAATCATGCCACATTAGCCGGGCATACATTCGAACATGAA
TTACGCATGGCAAGAGTACATGGTCTTCTTGGATCTGTTGATGCGAACCAGGGTCATCCTCTTTTAGGCTGG
GACACGGATGAATTTCCCACAGATTTATATTCTACGACATTAGCAATGTACGAAATCCTGCAAAATGGCGGC
CTTGGAAGCGGTGGCTTAAACTTTGACGCGAAGGTCAGAAGATCTTCTTTTGAGCCTGATGATTTAGTATAT
GCCCATATTGCAGGGATGGATGCATTTGCAAGAGGATTGAAAGTAGCCCACAAATTAATCGAAGATCGTGTG
TTTGAAGATGTGATTCAACATCGTTATCGCAGTTTTACTGAAGGAATTGGTCTTGAAATTACAGAAGGAAGA
GCTAATTTCCATACTCTTGAGCAATATGCGCTAAATAATAAAACAATTAAAAATGAATCTGGAAGACAGGAG
CGATTAAAACCTATATTGAACCAATAA

Streptomyces rubiginosus xylose isomerase

PROTEIN LENGTH = 388

MNYQPTPEDRFTFGLWTVGWQGRDPFGDATRRALDPVESVRRLAELGAHGVTFHDDDLIPFGSSDSEREEHV
KRFRQALDDTGMKVPMATTNLFTHPVFKDGGFTANDRDVRRYALRKTIRNIDLAVELGAETYVAWGGREGAE
SGGAKDVRDALDRMKEAFDLLGEYVTSQGYDIRFAIEPKPNEPRGDILLPTVGHALAFIERLERPELYGVNP
EVGHEQMAGLNFPHGIAQALWAGKLFHIDLNGQNGIKYDQDLRFGAGDLRAAFWLVDLLESAGYSGPRHFDF
KPPRTEDFDGVWASAAGCMRNYLILKERAAAFRADPEVQEALRASRLDELARPTAADGLQALLDDRSAFEEF
DVDAAAARGMAFERLDQLAMDHLLGARG*

DNA LENGTH = 1167 (Expression optimized)

ATGAATTATCAACCTACACCCGAAGACAGATTTACGTTTGGACTGTGGACGGTTGGATGGCAAGGTAGAGAT
CCCTTTGGAGACGCCACAAGAAGAGCTTTAGACCCAGTTGAGAGTGTAAGAAGATTGGCAGAATTAGGTGCA
CATGGTGTTACTTTTCATGACGACGACCTTATTCCATTCGGATCCAGTGATAGTGAAAGAGAAGAACATGTA
AAAAGATTTAGACAAGCTTTAGATGATACAGGTATGAAGGTTCCGATGGCTACTACAAATTTGTTTACTCAT
CCCGTATTTAAAGATGGTGGATTTACAGCTAACGACCGTGATGTTAGAAGATACGCTTTAAGGAAAACTATT
AGAAATATTGACTTAGCAGTTGAATTAGGTGCAGAAACTTACGTTGCCTGGGGCGGTAGAGAAGGTGCCGAA
TCTGGTGGGGCAAAAGATGTAAGAGATGCTTTGGATAGAATGAAAGAGGCCTTCGATTTATTGGGCGAATAT
GTTACCTCACAAGGCTATGATATTAGATTTGCAATCGAACCAAAACCAAATGAACCAAGAGGAGATATTTTA
CTACCGACAGTCGGACATGCATTAGCATTTATAGAGAGATTAGAAAGACCAGAACTTTATGGTGTAAATCCA
GAGGTAGGACACGAACAAATGGCAGGACTAAATTTTCCACATGGCATAGCACAAGCATTGTGGGCAGGCAAA
TTATTTCACATAGACTTAAACGGACAAAATGGTATCAAATATGACCAAGACTTAAGATTCGGAGCAGGAGAT
TTGAGAGCAGCATTTTGGTTGGTAGATTTATTAGAAAGTGCAGGTTATTCAGGCCCTAGACATTTTGATTTT
AAACCCCCAAGAACAGAAGATTTTGATGGAGTTTGGGCTTCAGCAGCAGGATGTATGAGAAATTATCTAATT
TTGAAAGAGAGAGCAGCAGCATTTAGAGCCGATCCAGAAGTTCAAGAAGCCTTAAGAGCAAGTAGACTAGAT
GAATTAGCAAGACCAACAGCAGCAGACGGATTACAAGCGTTATTGGATGATCGTTCAGCTTTCGAAGAATTC
GACGTAGACGCAGCGGCAGCAAGAGGTATGGCATTTGAAAGATTAGATCAATTAGCAATGGATCACTTACTA
GGAGCTAGGGGTTAA

Xanthomonas campestris xylose isomerase

PROTEIN LENGTH = 446

MSNTVFIGAKEYFPGIGKIGFEGRDSDNPLAFKVYDANKQVAGKTMAEHLRFAVAYWHSFCGNGADPFGPGT
RAYPWDVGNTALARAEAKSDAAFEFFTKLGVPYYCFHDIDLAPDADDIGEYENNLKHMVGIAKQRQADTGVK
LLWGTANLFSHPRYMNGASTNPDFNVVARAAVQVKAAIDATVELGGENYVFWGGREGYACLHNTQMKREQDN
MARFLTLARDYGRAIGFKGNFLIEPKPMEPMKHQYDFDSATVIGFLRQHGLDQDFKLNIEANHATLSGHSFE
HDLQVASDAGLLGSIDANRGNPQNGWDTDQFPTDLYDTVGAMLVVLRQGGLAPGGLNFDAKVRRESSDPQDL
FLAHIGGMDAFARGLEVADALLTSSPLETWRAQRYASFDSGAGADFANGTSTLADLATYAAGKGEPTQLSGR
QEAYENLINQYLTR*

DNA LENGTH = 1341 (Expression optimized)

ATGTCCAACACAGTTTTCATTGGTGCAAAAGAATATTTTCCGGGAATAGGTAAGATTGGATTCGAAGGTAGA
GACTCTGACAACCCTTTAGCATTTAAAGTTTATGACGCAAACAAACAAGTTGCGGGGAAAACTATGGCGGAA
CATTTAAGATTTGCAGTCGCATATTGGCATTCATTTTGTGGTAACGGTGCTGATCCATTTGGTCCAGGTACG
AGGGCTTACCCATGGGACGTTGGAAACACTGCACTTGCTAGAGCCGAGGCGAAATCTGATGCCGCTTTTGAA
TTTTTTACTAAATTAGGTGTACCATACTATTGCTTTCATGATATAGATTTAGCCCCAGATGCGGATGATATA
GGCGAATACGAAAATAATCTAAAACACATGGTTGGAATTGCAAAACAAAGACAAGCGGATACAGGTGTAAAA
TTATTATGGGGAACTGCTAACTTATTCTCACATCCAAGGTATATGAACGGTGCTAGCACTAATCCAGATTTT
AATGTTGTTGCAAGAGCTGCAGTACAAGTTAAAGCAGCCATAGATGCTACTGTTGAATTAGGAGGAGAGAAT
TACGTTTTTTGGGGTGGTAGGGAAGGATATGCATGTTTACACAATACACAAATGAAAAGAGAACAGGACAAT
ATGGCAAGGTTTTTAACATTAGCTAGAGATTATGGTAGAGCGATAGGATTTAAAGGTAACTTTTTGATAGAA
CCAAAACCGATGGAGCCTATGAAACATCAGTATGACTTCGATTCAGCTACTGTTATAGGATTTTTGAGACAG
CACGGCTTAGATCAAGATTTTAAATTAAACATAGAGGCAAATCACGCAACCCTATCAGGACATTCTTTTGAA
CATGATTTACAAGTTGCTTCAGATGCAGGGTTATTAGGATCAATTGACGCAAACAGAGGAAATCCACAAAAC
GGATGGGATACAGATCAATTTCCTACAGATTTATATGATACAGTTGGGGCGATGTTAGTAGTTTTAAGACAA
GGTGGACTAGCTCCCGGTGGTTTAAATTTCGACGCTAAAGTAAGAAGAGAAAGTAGCGATCCACAAGATTTG
TTTTTAGCACATATTGGCGGTATGGATGCTTTTGCAAGAGGATTAGAAGTAGCTGACGCTTTATTAACCAGT
TCACCATTAGAGACTTGGAGGGCACAAAGATACGCATCTTTTGATTCAGGCGCAGGAGCTGACTTTGCAAAT
GGCACTTCTACTCTAGCGGATCTAGCCACGTACGCTGCAGGAAAAGGAGAACCAACACAATTAAGCGGTAGA
CAAGAAGCATATGAAAATCTAATAAATCAATATTTGACCAGATAA

A.2. Putative xylose transporters

Putative transporters were selected using BLAST depending on sequence homology. The gene name, strain resource, Genbank accession number and primers used for cloning of each gene are listed in following table.

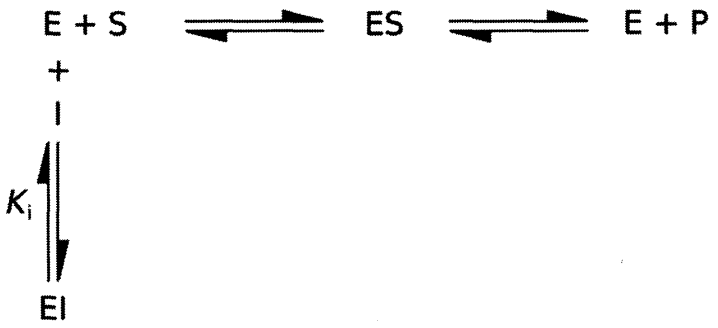
Table A-1. Putative xylose transporters

	Gene name	**Strain, Genbank Accession Number** **Primers used for cloning**	Restriction enzyme site
	PsXUT4	***Pichia stipitis*, XM_001386678**	
SUT1	P-PsXUT4-F	**ATTCTAGATGTCTTCGTTATTGACTAACG**	Xba I
	P-PsXUT4-R	**TTCTCGAGTTATTCCATCTCATTCAACTTGT**	Xho I
	Ps-SUT1	***Pichia stipitis*, XM_001387861**	
SUT2	P-PsSUT1-F	**TTACTAGTATGTCTTCTCAAGATATTCCTT**	Spe I
	P-PsSUT1-R	**TACTCGAGTTAAACATGTTCGTCAACA**	Xho I
	Ps-SUT2	***Pichia stipitis*, AF072808**	
SUT3	P-PsSUT2-F	**TTTCTAGATGTCCTCACAAGATTTACCCT**	Xba I
	P-PsSUT2-R	**TTCTCGAGCTAAACTTGCTCTTGCTCTT**	Xho I
	DhSUT1	***Debaryomyces hansenii*, CBS767, XM_458532**	
SUT4	P-DhSUT1-F	**AATCTAGATGGGTTTAGAAGATAATGCG**	Xba I
	P-DhSUT1-R	**TACTCGAGTTAGACTGAAGTGGTTTCAAT**	Xho I
	DhSUT2	***Debaryomyces hansenii*, CBS767, XM_457669**	
SUT5	P-DhSUT2-F	**AATCTAGATGACCAATACAGAACCAAATA**	Xba I
	P-DhSUT2-R	**TTCTCGAGTTACTCCTGTCTTTTATCTTCA**	Xho I
	DhSUT3	***Debaryomyces hansenii*, CBS767, XM_459306**	
SUT6	P-DhSUT3-F	**TATCTAGATGTTTAAGCAGGAGAATGGAA**	Xba I
	P-DhSUT3-R	**TACTCGAGCTATACTTGTTCATCAGAAGGT**	Xho I
	CaSUT1	**Candida albicans SC5314, XM_715292**	
SUT7	P-CaSUT1-F	**AATCTAGATGAGTGCAAATATCCAAGCT**	Xba I
	P-CaSUT1-R	**ATCTCGAGTTAAACGGAATTTTCATCAA**	Xho I
	CaSUT2	**Candida albicans SC5314, XM_710040**	
SUT8	P-CaSUT2-F	**TATCTAGATGTCAGAATTGACTGAACTT**	Xba I
	P-CaSUT2-R	**ATCTCGAGCTATTGTTGTAATAATCTACTT**	Xho I
	CgSUT1	**Candida glabrata CBS138, XM_444845**	
SUT9	P-CgSUT1-F	**ATACTAGTATGATCAGCGGCCCTAAGGAA**	Spe I
	P-CgSUT1-R	**TACTCGAGTCAATTTCTTGAGAATACTCT**	Xho I
	TrXLT1	**Trichoderma reesei, AY818402**	
SUT10	P-TrXLT1-F	**ACTCTAGATGTATCGGATTTGGAACATAT**	Xba I
	P-TrXLT1-R	**TACTCGAGTTAGACCTTTTCTTCGTGCT**	Xho I

	PsXUT5	***Pichia stipitis*, XM_001385925**	
SUT11	P-PsXUT5-F	**TAACTAGTATGACGGAAAGAAGCATTGG**	Spe I
	P-PsXUT5-R	**AACTCGAGTTACTTCTTTGTATTAACAAC**	Xho I
	ScHXT7	***S. cerevisiae*, Z31692**	
SUT12	ScHXT7-F	**TTACTAGTATGTCACAAGACGCTGCTATTG**	Spe I
	ScHXT7-R	**ATCTCGAGTTATTTGGTGCTGAACATTCT**	Xho I
	PsXUT6	***Pichia stipitis*, XM_001386552**	
SUT13	P-PsXUT6-F	**TTACTAGTATGTCCAGTGTTGAAAAAAGT**	Spe I
	P-PsXUT6-R	**TTCTCGAGTTAGCTGATGTTTTCGACAT**	Xho I
	ScHXT5	***S. cerevisiae*, X77961**	
SUT14	P-ScHXT5-F	**AAACTAGTATGTCGGAACTTGAAAACGC**	Spe I
	P-ScHXT5-R	**ATCTCGAGATTATTTTTCTTTAGTGAACAT**	Xho I
	Ps-XUT3	***Pichia stipitis*, XM_001387101**	
SUT15	P-PsXUT3-F	**AAATACTAGTATGAGAGAAGTTGGTATTCTTGATGTTG**	SpeI
	P-PsXUT3-R	**TTTTCTCGAGTTATTCTGACATTTCAATCGAGTTGC**	XhoI
	CaSUT3	***Candida albicans*, XM_714504**	
SUT16	P-CaSUT3-F	**AAAAACTAGTATGTCCTCAACTAATTCAACAG**	SpeI
	P-CaSUT3-R	**AATACTCGAGTTAAACGTGTTCCTCTGTTGG**	XhoI

Bold: restriction enzyme recognition sequence
Underline: start or stop codon

A.3. Xylitol inhibitory effect estimation



Assumption: competitive inhibition

$v_0=v_{max}\times S/(K_{m'}+S)$ where $K_{m'} = K_m\times(1 + I/K_i)$

From literatures:

K_m ~20-50mM

K_i ~5-15mM

From literatures and experiments:

S(xylose) ≈ 100mM

I (xylitol) = 1-10mM

Estimated results:

w/o inhibitor: $v_0 = 0.66v_{max}$

w/ inhibitor I = 1mM: $K_{m'}$ = 70mM; $v_{0'} = 0.59v_{max} = 0.88v_0$

w/ inhibitor I = 10mM: $K_{m'}$ =150mM; $v_{0'} = 0.4v_{max} = 0.6v_0$